INFLUENCES OF SENSORY MODALITY AND STIMULUS TYPE ON PROCESSING TASK-IRRELEVANT STIMULI

A DISSERTATION TO BE SUBMITTED TO THE GRADUATE
DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF
THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PSYCHOLOGY

MAY 2019

By

Maegen E. Walker

Dissertation Committee:

Scott Sinnett, Chair Person
Patricia Couvillon
Therese Grüter
Kalpana Kallianpur
Grayden Solman
Jonas Vibell

Keywords: cognition, attention, facilitation, multisensory processing, task-irrelevant stimuli

ACKNOWLEDGEMENTS

The research conducted for this dissertation would not have been possible without the continued support and guidance of my advisor, Dr. Scott Sinnett, and help from the exemplary research assistants that have volunteered their time in the Perception and Attention Research (PAR) Laboratory at the University of Hawaiʻi at Mānoa. I would also like to thank my parents, Allyson Williams and Jack Koegle, for cheering me on throughout my education and professional pursuits. Finally, a special ‘thank you’ to my husband, Thomas Walker, for his unyielding patience and understanding; being my sounding board, voice of reason, and biggest advocate.

Thank you!

ABSTRACT

The processing of unattended, task-irrelevant, stimuli that are frequently presented in temporal-alignment with an attended target (i.e., target-aligned or TA) in an attention-demanding task is often facilitated. This facilitation results in higher recognition rates for TA items compared to other unattended items, presented with equal frequency, that do not appear in temporal-alignment with an attended target (i.e., non-aligned or NA), when encountered later during a surprise recognition test. Previous investigations exploring the facilitated processing for unattended TA items have traditionally focused on word processing in the visual modality. Relatively few studies have explicitly examined the role of sensory modality or the types of stimuli being presented when evaluating the extent to which unattended information may be facilitated in an attention-demanding task. Because humans live in a multisensory environment in which we are exposed to a variety of stimuli, it is important to investigate the extent to which processing for different types of items may be facilitated when unattended and whether or not the sensory modality in which these stimuli are presented plays a role in facilitation. Arguably, vision and audition are the two most dominant sensory modalities for humans, and two main forms of information that are often encountered in daily life include lexical (i.e., words, both written and spoken) and non-lexical information (i.e., pictures and sounds). The experiments presented in this doctoral dissertation explore the role that sensory modality (vision and audition) and stimulus type (lexical and non-lexical) have in the facilitated processing of unattended, task-irrelevant stimuli.

Experiment 1 focused on comparing facilitation rates for lexical and non-lexical stimuli under unimodal visual conditions (Experiment 1a) and unimodal auditory conditions (Experiment 1b). It was hypothesized that all TA items would undergo facilitated processing leading to higher recognition rates compared to NA items and that non-lexical information would be preferentially facilitated, in general, compared to lexical information. Collectively, the results of Experiment 1 demonstrate that under unimodal visual conditions, the facilitatory

effects of target-alignment remain robust, with higher recognition rates for TA items compared to NA items observed regardless of stimulus type. Additionally, unattended non-lexical items (i.e., pictures) appear to be processed more extensively (i.e., facilitated), leading to higher recognition rates overall, compared to lexical items (i.e., written words), and the impact of target-alignment remains uniform across these stimulus dimensions. However, when presented under unimodal auditory conditions (Experiment 1b), target-alignment does not appear to play a critical role in facilitation as TA items were not recognized more often when compared to NA items. Additionally, there was no significant difference in recognition rates for unattended non-lexical items (i.e., sounds) when compared to lexical items (i.e., auditory words), suggesting that stimulus type also has little bearing on the facilitated processing of auditory items that were unattended and task-irrelevant.

Experiment 2 extended the unimodal stimulus presentation conditions from the first experiment to cross-modal conditions, while comparing facilitation rates for lexical and non-lexical stimuli. There were two conditions; an auditory/visual condition in which the unattended dimension was presented in the visual modality while performing an attention-demanding auditory task (Experiment 2a), and a visual/auditory condition in which the unattended dimension was presented in the auditory modality while performing an attention-demanding visual task (Experiment 2b). As with Experiment 1, a main effect for target-alignment was anticipated, as was a main effect for stimulus type, under both conditions (auditory/visual and visual/auditory). Results from Experiment 2 mirrored those of Experiment 1. In Experiment 2a, the effects of target-alignment and stimulus type remained robust for unattended visual information when presented concurrently with an attention-demanding auditory task. Specifically, TA items were recognized more often compared to NA items and non-lexical items (i.e., pictures) were recognized more often when compared to lexical items (i.e., written words). In Experiment 2b, the effects of target-alignment and stimulus type remained inconsequential. TA items were not recognized more often than NA items and there was no difference in recognition rates between lexical (i.e., auditory words) and non-lexical items (i.e., sounds).

Taken together, Experiments 1 and 2 demonstrate that facilitated processing for unattended lexical and non-lexical information may proceed differently depending on the sensory modality in which those items are presented. In the visual modality, target-alignment facilitates processing for unattended items, regardless of stimulus type, and, regardless of target-alignment, unattended pictures appear to be processed more extensively than unattended written words. In the auditory modality, target-alignment does not appear to play a critical role in the processing of unattended information, as there was no difference in recognition rates between TA and NA items. Furthermore, there appears to be no difference in the extent to which unattended auditory words and sounds are processed. The underpinning theoretical rationale for the divergent patterns in facilitation rates of lexical and non-lexical stimuli between the two sensory modalities are extensively explored in this dissertation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS .... **i**
ABSTRACT .... **ii**
LIST OF TABLES .... **viii**
LIST OF FIGURES .... **ix**

1. General Introduction **1**

2. Information Processing Approach **3**
2.1. Model of Information Processing 4
2.2. General Principles of an Information Processing Model 5

3. Attention and Information Processing **8**
3.1. Attention as a Gating Mechanism 8
3.2. Attention and Inhibition 12
3.3. Attention and Facilitation 15
3.3.1. Facilitation for Explicitly Ignored Information 19

4. Multisensory Integration **24**
4.1. Interactions Between Visual and Auditory Senses 24
4.2. Interactions Between Multisensory Integration and Attention 27
4.2.1. Cross-Modal Spreading 27
4.2.2. Cross-Modal Cueing 30
4.2.3. Cross-Modal Prior Entry 32

5. Attention and Processing Lexical and Non-Lexical Information **36**
5.1. Attended Lexical and Non-Lexical Information 36
5.1.1. Visual Modality 36
5.1.2. Auditory Modality 38
5.2. Ignored Lexical and Non-Lexical Information 40
5.2.1. Visual Modality 40
5.2.2. Auditory Modality 43

6. Research Direction and Aims **48**
6.1. Aims of this Dissertation 49

7. Experiment 1: Unimodal Conditions **51**
7.1. Experiment 1a: Unimodal Visual Conditions 51
7.1.1. Participants 51
7.1.2. Stimuli 52
7.1.2.1. Attended Written Words / Ignored Pictures 53
7.1.2.2. Attended Pictures / Ignored Written Words 54
7.1.2.3. Surprise Recognition Test 54
7.1.3. Procedure 55
7.1.4. Predictions 56
7.1.5. Results 58
7.1.5.1. Primary Task 58
7.1.5.2. Surprise Recognition Test 60
7.1.5.2.1. Ignored Written Words 65
7.1.5.2.2. Ignored Pictures 66

7.1.5.2.3. Analysis by Target-Alignment 67
7.1.6. Discussion: Experiment 1a 68

7.2. Experiment 1b: Unimodal Auditory Conditions 72
7.2.1. Participants 73
7.2.2. Stimuli 74
7.2.2.1. Attended Auditory Words / Ignored Sounds 74
7.2.2.2. Attended Sounds / Ignored Auditory Words 74
7.2.2.3. Surprise Recognition Test 75
7.2.3. Procedure 75
7.2.4. Predictions 76
7.2.5. Results 79
7.2.5.1. Primary Task 79
7.2.5.2. Surprise Recognition Test 81
7.2.5.2.1. Ignored Auditory Words 84
7.2.5.2.2. Ignored Sounds 85
7.2.5.2.3. Analysis by Target-Alignment 86
7.2.6. *Discussion: Experiment 1b* 86

8. Experiment 2: Cross-Modal Conditions 92
8.1. Experiment 2a: Cross-Modal Ignored Visual Conditions 92
8.1.1. Participants 92
8.1.2. Stimuli 93
8.1.2.1. Attended Sounds / Ignored Written Words 93
8.1.2.2. Attended Auditory Words / Ignored Pictures 94
8.1.2.3. Surprise Recognition Test 94
8.1.3. Procedure 94
8.1.4. Predictions 95
8.1.5. Results 97
8.1.5.1. Primary Task 97
8.1.5.2. Surprise Recognition Test 100
8.1.5.2.1. Ignored Written Words 104
8.1.5.2.2. Ignored Pictures 105
8.1.5.2.3. Analysis by Target-Alignment 106
8.1.6. Discussion: Experiment 2a 107

8.2. Experiment 2b: Cross-Modal Ignored Auditory Conditions 111
8.2.1. Participants 112
8.2.2. Stimuli 112
8.2.2.1. Attended Pictures / Ignored Auditory Words 113
8.2.2.2. Attended Written Words / Ignored Sounds 113
8.2.2.3. Surprise Recognition Test 113
8.2.3. Procedure 113
8.2.4. Predictions 114
8.2.5. Results 116
8.2.5.1. Primary Task 116
8.2.5.2. Surprise Recognition Test 118
8.2.5.2.1. Ignored Auditory Words 121

8.2.5.2.2. Ignored Sounds 122
8.2.5.2.3. Analysis by Target-Alignment 123
8.2.6. Discussion: Experiment 2b 124
9. Discussion, Limitations, and Future Directions 128
9.1. Discussion 128
9.1.1. Experiment 1a: Purpose and Key Findings 130
9.1.2. Experiment 1b: Purpose and Key Findings 131
9.1.3. Experiment 2a: Purpose and Key Findings 132
9.1.4. Experiment 2b: Purpose and Key Findings 134
9.1.5. General Discussion 135
9.2. Limitations 147
9.3. Future Directions 151
Appendix A: IRB Approval Letter 154
Appendix B: Consent Form 156
Appendix C: Debriefing Form 159
Appendix D: List of Stimuli 160
***References* 166**

LIST OF TABLES

Table 1. A Summary of Results from the Critical Analyses 130

LIST OF FIGURES

Figure 1. Flow Chart of Information Processing and Memory Storage 4

Figure 2. Example of Attentional Blink Task 11

Figure 3. Schematic of Cueing Paradigm 13

Figure 4. Diagram Events in Stelmach and Herdman (1991) 16

Figure 5. Example of Array Configurations in Ahissar and Hochstein (1993) 18

Figure 6. Example of RSVP task and DRD in Watanabe et al. (2001) 21

Figure 7. Example of Stimuli and Task in Busse et al. (2006). 29

Figure 8. Diagram of Stimuli in Spence and Driver (1996) 31

Figure 9. Schematic of Primary Task in E1a 56

Figure 10. D' Rates During the Primary Task in E1a 60

Figure 11. Accuracy Scores for the Surprise Recognition Test in E1a 62

Figure 12. D' Rates During the Surprise Recognition Test in E1a 63

Figure 13. Results from the Critical Analysis of the Surprise Recognition Test in E1a 65

Figure 14. Recognition Rates for NA Items During the Surprise Recognition Test in E1a 68

Figure 15. Schematic of Primary Task in E1b 76

Figure 16. D' Rates During the Primary Task in E1b 80

Figure 17. Accuracy Scores for the Surprise Recognition test in E1b 82

Figure 18. Results from the Critical Analysis of the Surprise Recognition Test in E1b 84

Figure 19. Schematic of Primary Task in E2a 95

Figure 20. D' Rates During the Primary task in E2a 98

Figure 21. RT to Identified Targets During the Primary Task in E2a 99

Figure 22. Accuracy Scores for the Surprise Recognition Test in E2a 101

Figure 23. D' rates During the Surprise Recognition Test in E2a 102

Figure 24. Results from the Critical Analysis of the Surprise Recognition Test in E2a 104

Figure 25. Recognition Rates for NA Items During the Surprise Recognition Test in E2a 107

Figure 26. Schematic of Primary Task in E2b 114

Figure 27. D' Rates During the Primary Task in E2b 117

Figure 28. Accuracy Scores for the Surprise Recognition Test in E2b 119

Figure 29. Results from the Critical Analysis of the Surprise Recognition Test in E2b 121

Figure 30. Recognition Rates for NA Items During the Surprise Recognition Test in E2b 124

Figure 31. Results from the Critical Analyses on the Surprise Recognition Test for all Experiments 136

1. General Introduction

In order to complete any given task, such as driving a car or searching for an item at a grocery store, attention must be directed toward task-relevant stimuli (i.e., information or items necessary for completing the task at hand), while simultaneously avoiding the allocation of attentional resources toward task-irrelevant stimuli (i.e., additional available information or items that are not related to the task). For example, in order to successfully navigate a vehicle, we must direct our attention toward relevant information such as road signs and traffic signals, while ignoring task-irrelevant information such as an incoming text message, a distracting billboard, or our favorite song on the radio. While an ongoing debate regarding the extent to which unattended information may be processed exists, it is well established that ignored information *can* be processed and potentially modulate human behavior (Dehaene et al., 1998; Dewald, Sinnett, & Doumas, 2013; Rees, Russell, Frith, & Driver, 1999; Seitz & Watanabe, 2003, 2005; Sinnett, Costa, & Soto-Faraco, 2006; Tsushima, Sasaki, & Watanabe, 2006; Tsushima, Seitz, & Watanabe, 2008; Watanabe, Náñez, & Sasaki, 2001).

From an information-processing perspective (Barber, 2015; Lachman, Lachman, & Butterfield, 2015; Simon, 1979), we assume that information that is directly attended to is processed more quickly or extensively (i.e., facilitated – resulting in a more robust mental representation and/or more fluent responses to that information) when compared to information that is not attended to, and is therefore more likely to be encoded, perceived, stored in long-term memory, and subsequently recognized later. However, the extent to which ignored information may be facilitated is likely to be subject to a number of constraints. For example, a growing body of literature has outlined the role of temporal presentation in the facilitation of task-irrelevant stimuli (Dewald et al., 2013; Seitz & Watanabe, 2003, 2005; Swallow & Jiang, 2010, 2011; Watanabe et al., 2001; Watanabe & Sasaki, 2015), but additional factors have yet to be established. Two likely candidates may be whether a single sensory modality is used to process incoming information, or if instead multiple sensory modalities are involved, and the type of stimulus that is being processed (i.e., lexical vs. non-lexical information).

Early investigations have demonstrated that information processing for attended information may be modulated by the number of sensory modalities involved in the task (Colavita, 1974; Freides, 1974; Posner, Nissen, & Klein, 1976), and the type of stimuli presented to participants (Amit, Algom, & Trope, 2009; Kensinger & Schacter, 2006; Kieras, 1978; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996). Therefore, it is reasonable to assume that these two factors may also modulate the extent to which ignored, task-irrelevant, information may be facilitated. This doctoral dissertation focuses on exploring the role of sensory modality and stimulus type in the facilitated processing for unattended (i.e., ignored), task-irrelevant, information.

In order to elucidate how sensory modality and stimulus type may affect the processing of unattended, task-irrelevant information, it is important to describe the information-processing approach to studying human cognition, giving special consideration to the role that attention has in such a system. This will be followed by an exploration of the relevant mechanisms of attention and how they may contribute to the processing of ignored information, after which a brief discussion regarding how this may be influenced by sensory modality and stimulus type will be offered. Finally, focus will be turned to the individual experiments, results, discussions, future directions, and limitations.

2. Information Processing Approach

The information processing approach to studying human cognition arose from the cognitive revolution of the 1950's and 60's. As a result, researchers began to draw comparisons between the computational algorithms of computer science and human cognitive capabilities. Rather than simply responding to stimuli, information processing assumes a step-wise series of events in which incoming information is analyzed for meaning, organized according to a set of rules or algorithms, stored in long term memory, and used in decision making and response selection processes (Barber, 2015; Lachman et al., 2015). In line with this approach, Shiffrin and Atkinson (1969) described a model of information processing within the human memory system outlining a series of processing stages (Figure 1). While the model has been updated to account for some of the dynamic aspects of human working memory (previously referred to as short term memory; see Baddeley, 1992) the foundation of the information processing system remains largely similar today.

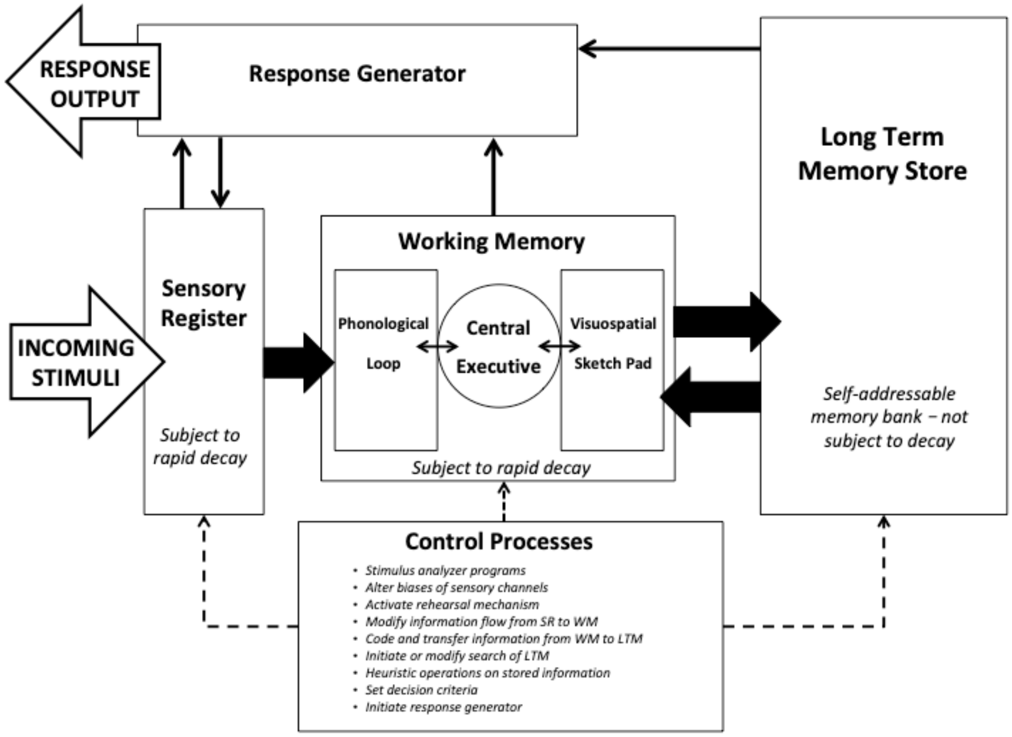


Figure 1. Modified flow chart (short term memory component has been updated to reflect the more modern conceptualization of working memory, see Baddeley, 1992) of the stages involved in information processing and memory storage outlined by Shiffrin and Atkinson (1969). Solid lines indicate information transfer between contiguous stages; dashed lines indicate connection pathways between the different stages – allowing for information arrays residing in different parts of the system to be compared in parallel as well as the transmission of control signals mediating information transfer, rehearsal mechanisms, decision criteria, and response selection.

2.1. Model of Information Processing

According to the information processing approach (Atkinson & Shiffrin, 1968, 1971; Baddeley, 1992, 2012; Baddeley & Hitch, 1974; Hitch & Baddeley, 2017; Huitt, 2003; Shiffrin & Atkinson, 1969; Skóra & Wierzchoń, 2016; Sørensen, 2017), incoming information received at the senses is processed first by a sensory register, which is a transient storage system maintaining sensory information in its original form for a fraction of a second (e.g., less than a second in the case of visual sensory memory, i.e., iconic memory). The selection of relevant information begins in this stage, while extraneous information may be subject to rapid decay and subsequently forgotten. Selected information may be transferred to a short-term store, more recently conceptualized as working memory (see Baddeley, 1992), where it may be processed further (i.e., evaluated for meaning, related to existing information, cognitively manipulated, etc.). From here, decision criteria may be applied followed by immediate response selection and execution; the information may, again, decay and be forgotten, or the information may be transferred to the long-term memory store where it is incorporated into existing knowledge and can be retrieved for use at a later time. Thus, information that is "passed forward" along the successive stages is thought to receive more extensive processing and evaluation compared to information that is filtered out at an earlier stage, which is assumed to undergo some form of decay. Furthermore, information that may otherwise be lost (i.e., forgotten) may be carried forward along these stages by underlying mechanisms that promote (i.e., facilitate) further processing and subsequent storage. This idea will be explored in greater detail later.

2.2. General Principles of an Information Processing Model

It should be noted that while the information processing model discussed here (Shiffrin & Atkinson, 1969) is among the most widely known models of the human memory system today, additional models have been proposed with variations on the idea that information may be processed in stages of some sort (Baddeley, 1992; Bransford, 1979; Craik & Lockhart, 1972; McClelland, 2000; Rumelhart, McClelland, & PDP Research Group, 1987). Despite different conceptualizations of *how* information may be processed, a broad set of principles may generally be agreed upon that outline constraints within each of the proposed models. These constraints are 1) a limited capacity system, 2) two-way flow of information, and 3) a control mechanism.

The first principle assumes that our cognitive system operates with a limited capacity. This means that the amount of information that can be processed by the system at any given time is subject to a series of constraints that likely occur at specific points within the system (Huitt, 2003). Each model posits the existence of various gating mechanisms and bottlenecks designed to limit the flow of incoming information in order to avoid cognitive overload. For example, when examining visual information processing in humans, Sperling asserted that memory "sets a limit on a process that is otherwise rich in available information" (1960, p. 26) suggesting that memory capacity, itself, may serve as a type of gating mechanism. Indeed, more recent investigators have also posited the existence of bottlenecks and gating mechanisms controlling the flow of incoming information, thereby allowing the system to process the selected or task-relevant information and respond accordingly (Broadbent, 1957; Duncan, 1980; Duncan, Martens, & Ward, 1997; Lavie, 1995, 2010; Treisman, 1969; Treisman & Gelade, 1980; Wolfe, 2001; Wolfe & Gray, 2007).

The second principle, two-way flow of information, assumes that the flow of information being processed by our cognitive system is bi-directional. Incoming sensory information from the environment is compared to information already stored in memory. Thus, we compare our current experience to previous experiences and use this pre-existing information as a guiding mechanism to organize, categorize, and generally make sense of the sensory stimuli in our environment (Huitt, 2003). This process may also help dictate which pieces of information are selected for further processing by the system in a form of top-down cognitive control.

The third principle, a control mechanism, asserts that some form of central executive control (Baddeley, 1992) mediates the flow of information through the system – both incoming and outgoing. Specifically, the central executive controls information encoding, transfer, processing, storage, retrieval, and utilization (Huitt, 2003). It should also be noted that the central executive is not separate from the cognitive system, but rather is integrated into it. This means that the central executive also requires some of the limited available processing capacity. Thus, when one is engaged in a difficult task the central executive utilizes additional resources in order to successfully complete that activity, leaving fewer resources available to process the incoming information, adding an additional level of constraint.

While a general set of operating principles in an information processing system may be broadly agreed upon, the precise mechanisms that may underlie these operations are subject to debate. While there may well be many mechanisms, working in tandem, to help our cognitive system select, organize, interpret, store, and utilize information, attention may be one that has a large influence over the extent to which a piece of information may be processed and subsequently stored in memory. The following chapter will explore the role of attention in an information processing system in more detail.

3. Attention and Information Processing

Despite divergent models of how our cognitive system may process information, it is widely accepted that attention plays a pivotal role in aiding the selection of relevant items, combining features together into perceptual wholes, and filtering irrelevant information (Baddeley & Weiskrantz, 1993; Broadbent, 1954; Lachter, Forster, & Ruthruff, 2004; Posner & Cohen, 1984; Posner, Snyder, & Davidson, 1980; Stelmach & Herdman, 1991; Treisman, 1969; Treisman & Gelade, 1980; Wolfe, 2001; Wolfe & Gray, 2007). Thus, the extent to which available information is "passed forward" to later processing stages is likely to be mediated, in part, by attentional limitations. While a comprehensive review of attention and its mechanisms is beyond the scope of this dissertation, the current chapter will explore some of the functions of attention in an information processing model in detail. Specifically, consideration will be given to the role of attention as a potential gating mechanism, inhibiting and facilitating the flow of attended information, and finally facilitating processing of explicitly ignored information.

3.1. Attention as a Gating Mechanism

The previous chapter outlined several principles of a general model of information processing, in particular, the idea that cognitive resources related to information processing are capacity limited. Therefore, in order to control the flow of incoming information, some form of gating mechanism has been proposed to keep the system from overload when faced with the seemingly limitless amount of information that is available to us. While it is likely that there are many potential mechanisms in place designed to help control the flow of information, attention has historically been thought of as playing a central role in this process. The exact manner in which attention helps control the flow of information is not well agreed upon and, as a result, several conceptualizations of this process have been proposed. Some of the earliest models, such as Broadbent's Filter Theory (Broadbent, 1958), Treisman's Attenuation Theory (Treisman, 1960, 1969), and Treisman's Feature Integration Model (Treisman & Gelade, 1980), distinguish between a pre-attentive stage and an attentive stage in information processing that are separated by a "filter." In the pre-attentive stage, physical properties, such as pitch and sound location, or shape and color, are extracted from all incoming stimuli in a *parallel* manner – meaning that processing for these individual features happens simultaneously – and that this process is relatively un-guided by top-down control mechanisms such as strategy or goal oriented intents of the agent. Items that are "selected" (Theeuwes, 1993) in the pre-attentive stage are allowed to pass through the filter and enter the next stage where the information is further evaluated and manipulated within the system. From this perspective, the process of selection occurs when attention is *applied* to the stimulus, thereby forwarding the information in the system, where more complex evaluations and manipulations of the information take place.

Thus, information that is "selected" may pass through the filter from the pre-attentive stage and is thought to enter the attentive stage of processing. Here, complex psychological properties such as stimulus identity or meaning are extracted from the selected stream of information allowing for the formulation of specific response operations. This stage is generally assumed to be capacity limited, both in the *amount* of information that can be processed at a time and by the spatial *location* that the attended information comes from, leading to serial information processing (i.e., one-by-one). This stage is also generally thought to be subjected to top-down guidance control wherein observer influences, such as expectations and task-related objectives, help to further guide the deployment of attention among selected items (Duncan, 1980; Humphreys, 1981; Posner, 1980; Theeuwes, 1993; Treisman & Gelade, 1980). In this way, attention may also be thought of as a mechanism that *guides* the selection of potentially relevant information. This means that once attention has been deployed to a particular location or item, this piece of information has been "selected" for further processing and is then allowed to pass through the filter and advance along the system. After this occurs, a series of constraints are applied to the incoming information, which determine where in space additional items of relevance may be selected from, what features of the selected items are highlighted (or focused on), and the order in which those items are processed. Indeed, Treisman and Gelade (1980) hypothesized that attention operates as the "glue" that helps bind individual features into perceivable objects between the pre-attentive stage and the attentive stage.

Other models posit more than one filter in the system and that the flow of information through those filters may be mediated, at least in part, by attentional processes. One such model (Wolfe, 2001; Wolfe, Cave, & Franzel, 1989; Wolfe & Gray, 2007), asserts that early vision occurs in a series of parallel processes that are subjected to a mandatory selective bottleneck – akin to the pre-attentive stage of other models. Those features that are passed through the bottleneck form a "guiding representation" that provides the input for later object recognition processes (Wolfe & Gray, 2007). Information that is allowed to pass through the first bottleneck, as well as some information that may bypass the bottleneck altogether, is subjected to a second bottleneck that limits performance on decision-making and response criteria. Top-down guidance, such as task constraints or viewer expectations, also operates in a feedback manner to help shape the guiding representation. In this model, attention is thought to function as the *control* exerted over the selection of items to be passed through the successive bottlenecks. Thus, an initial early filtering mechanism is responsible for item recognition, while a second filtering mechanism further prunes the incoming information for those items that will be acted upon; attentional processes may help guide the selection of items to be passed through each successive bottleneck.

Wolfe and Gray (2007) argue that the idea of two bottlenecks – one for item selection and one for response selection – is supported by behavioral evidence from the attentional blink (AB) paradigm (Shapiro, 1994). AB occurs when participants fail to report the occurrence of a second target (T2) presented in a rapid serial visual presentation (RSVP) of items if it appears within 200-500ms after the first target (T1) in the stream (Figure 2). However, if T2 appears either immediately after T1 (i.e., <200ms) or after a delay longer than 500ms, it is more easily identified (Chun & Potter, 1995; Shapiro, 1994).

(a) (b)

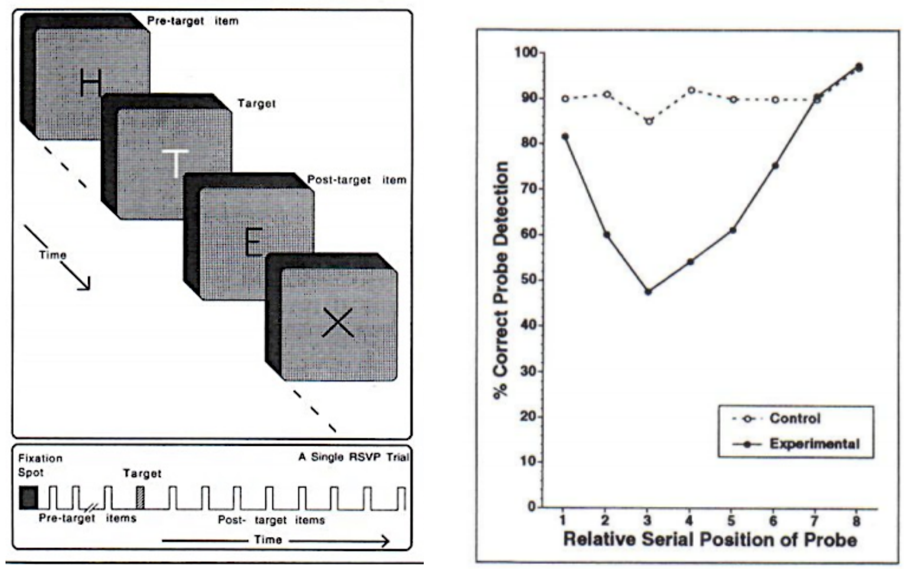


Figure 2. (a) Example of the AB task taken from Shapiro (1994). The white ‘T’ represents T1 while the black ‘X’ represents T2. (b) Mean percentage of trials wherein the T2 was successfully identified, plotted as a function of serial position in the RSVP stream after the appearance of T1. Open circles represent the control condition, where participants were asked to ignore T1 and just identify T2. Closed circles represent the experimental condition, where participants were instructed to identify both T1 and T2. Note the drop in performance for T2 identification when presented within 200-500ms after T1 (i.e., serial positions 2 – 5), compared to shorter delays (i.e., serial position 1), followed by an increase in T2 recognition for delays longer than 500ms (i.e., serial positions 6 – 8).

From an information processing perspective, once T1 is selected, it passes through the first bottleneck and target identification occurs, which takes about 200-500ms (Wolfe & Gray, 2007). When T2 is presented within this time frame, it may also pass through the first filter but become blocked by the second filter because attention is still diverted to T1 and decision processes or response criteria are beginning to form. When T2 is presented within a very short window after T1, both items pass through both filters via parallel processing, and when T2 is presented long after T1 (i.e., >500ms) both items are free to pass through the successive bottlenecks in a serial manner. In this perspective, attention is thought to work in conjunction with temporal presentation to mediate the extent to which one or more pieces of information are able to pass through each successive filter.

This section has outlined the role of attention as a gating mechanism designed to help guide selection processes for items of relevance. A consequence of this process is that some information may either be inhibited (i.e., suppressed from further processing and subsequently forgotten - or not perceived at all) or it may be facilitated (i.e., passed forward along the system, analyzed for meaning, and become capable of influencing behavior). The next section of this chapter will explore these concepts in more detail.

3.2. Attention and Inhibition

Despite the fact that our attention can be involuntarily captured by some environmental stimuli (Anderson, Laurent, & Yantis, 2011; Yantis, 1996; Yantis & Hillstrom, 1994), in order to focus our attention toward task-relevant information we must be able to inhibit (i.e., suppress), at least in part, orienting to irrelevant, or previously attended, spatial locations and/or irrelevant information occurring at a spatial location to which attention is already focused. Posner and Cohen (1984, see also Dodd & Pratt, 2007) provided an example of the former condition when they demonstrated that initial orienting to one spatial location is later accompanied by an inhibited capacity to re-orient to that previously attended spatial location under certain laboratory conditions. This processes is known as "inhibition of return" (IOR). The IOR phenomenon was demonstrated using a classic cueing paradigm (Posner, 1980, see Figure 3), wherein participants viewed a display containing three boxes arranged parallel to one another. Participants maintain fixation on the center box while an exogenous cue (i.e., external cue – in this case, a transient increase in luminance) would be presented around either the left or right peripheral box. This exogenous cue was designed to draw participant's covert attention (i.e., attentional shift without eye movement) toward that spatial location. After a variable delay (i.e., stimulus onset asynchrony or SOA), a target black square appeared in one of the three boxes. The location of the black square could either be congruent with the cue (i.e., one the same side), incongruent (i.e., on the opposite side), or neutral (i.e., in the center box, Figure 3).

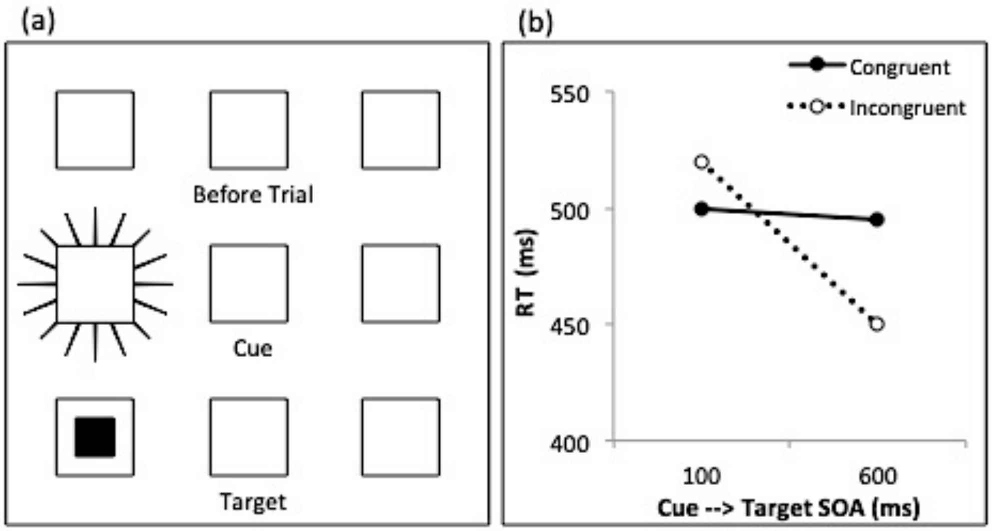


Figure 3. (a) Schematic representation of the traditional cueing paradigm used in Posner and Cohen (1984) depicting a congruent trial (which produces IOR under certain circumstances). Here, attention is drawn to a peripheral box with an exogenous cue. After a variable delay a target could appear in any of the three boxes. (b) Recreated graph depicting the cuing effect for short SOAs and the emergence of IOR over time (longer RTs for congruent trials (solid line) compared to incongruent trials (dotted line). If the target appeared in the same box as the cue (i.e., a congruent trial) within a very brief time window (i.e., between 500 and 1000ms) participants were slower to respond compared to when the target appeared in the box on the opposite side (i.e., an incongruent trial). The observed slower RT on congruent trials with an SOA between 500 and 1000ms is interpreted as IOR.

IOR was demonstrated by comparing participant reaction times (RTs) to targets appearing on the congruent side with those of the incongruent side. When the SOA between the exogenous cue and the target was greater than 300ms, RTs to targets appearing on the congruent side were significantly slower than RTs to targets appearing on the incongruent side. Furthermore, this delay in responding to congruent targets could last up to 1.5 seconds after the cue had been presented (Posner & Cohen, 1984, see also Dodd & Pratt, 2007). The authors concluded that IOR occurs under conditions in which attention is *summoned*, via an exogenous cue, to a particular location and subsequently diverted away again. IOR also appears to occur in cases of both covert (*without* eye movement) and overt (*with* eye movement) attentional shifts. Thus, IOR may be stimulus driven, spatially locked, and serves to maximize sampling of the visual field. Once attention is diverted away from a cued location, processing of information presented at that spatial location is temporarily inhibited compared to information presented at other, not-yet-sampled, locations. In this way, IOR prevents attention from being re-captured by events occurring at a location to which attention has recently been deployed thereby biasing search toward locations in which a target has not yet been detected. The authors concluded that this inhibitory mechanism plays an important role in directing future attentional acts (Posner & Cohen, 1984)[1].

Deploying attention to a specific set of items within a particular spatial location can also result in inhibited processing for task-irrelevant information co-occurring in that same region of space. In some cases, we fail to perceive highly salient information in our visual field because our attention is otherwise focused. This phenomenon is referred to as 'inattentional blindness' (Mack & Rock, 1998a, 1998b). The paradigm used to demonstrate this phenomenon in a natural setting was initially developed by Neisser and Becklen (1975), but was later popularized by Simons and Chabris (1999). In their experiment (Simons & Chabris, 1999), participants watched a video of a complex dynamic scene in which there were two small groups of individuals (i.e., teams), one wearing white shirts and one wearing black shirts. Each team would pass a basketball to their teammates while ducking and weaving around the other people nearby. Participants were instructed to focus their attention on one team (either black or white players) and count how many times the players on that team

[1] It should be noted that more recent investigations suggest that IOR is a function of attention in search behavior specifically, which may not apply under conditions of passive or free viewing of a visual scene (Dodd, Van der Stigchel, & Hollingworth, 2009) and may even be overridden by facilitation of return, a pattern of eye saccades wherein there is a tendency to voluntarily return to the last fixation point in a 1-back, sudden onset, scene viewing task (FOR, Smith & Henderson, 2009).

passed their basketball to their teammates. Approximately mid-way through the video a person wearing a gorilla suit walked through the players, thumped their chest, and exited the scene – the gorilla was in plain view for about five seconds.

After viewing the video, and counting ball passes, participants were asked how many passes the team they were monitoring made to ensure attention was focused on the task. They were also asked a series of questions, with escalating specificity, regarding what else they might have perceived, in an effort to determine if they noticed the gorilla. Despite the gorilla appearing on screen for an extended duration (approximately five seconds), and being clearly visible during that time, 46% of all viewers failed to perceive it (Simons & Chabris, 1999). These findings imply that the ability to perceive highly salient visual information presented in an attended location is inhibited when attention is diverted to other stimuli in that same area. It should be noted that the extent to which unattended information is inhibited is a topic of ongoing debate; indeed, more recent studies suggest that, even when actively ignored, information can still be processed and perceived to some extent (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe et al., 2001; Walker, Ciraolo, Dewald, & Sinnett, 2017). This seemingly paradoxical finding will be discussed in detail in the following sections.

3.3. Attention and Facilitation

Focusing attention away from a stimulus may serve to decrease the likelihood or speed at which that information is processed. On the other hand, processing for attended information tends to be facilitated or enhanced in some way. This means that information to which attention is focused is more likely to be detected and responses to those stimuli tend to be faster and more accurate (Ahissar & Hochstein, 1993; Liu, Abrams, & Carrasco, 2009; Posner, 1980; Stelmach & Herdman, 1991; Treisman & Gelade, 1980). For example, Stelmach and Herdman (1991) examined the relationship between attentional allocation and processing speed for visual stimuli with the use of a temporal order judgment (TOJ) task. In a TOJ task, participants are presented with two stimuli appearing at various SOAs, and are asked to determine which of the two stimuli appeared first. These authors suggested that directed attention facilitates the speed with which information is processed (see also Mcdonald, Teder-Sälejärvi, Russo, & Hillyard, 2003; Shore, Spence, & Klein, 2001; Spence, Shore, & Klein, 2001). In a modified cueing paradigm (see Posner, 1980 for an example of a cueing paradigm), participants in Stelmach and Herdman's experiment viewed an array of dots arranged in to three squares – one at fixation and two on each side – and were asked to attend either to the center, right, or left square. An additional dot would then transiently appear for 10ms in either the left or right square in the array followed by a very short, variable, SOA (i.e., 0 – 70ms) before the presentation of another transient dot for an additional 10ms in the square on the opposite side (Figure 4).

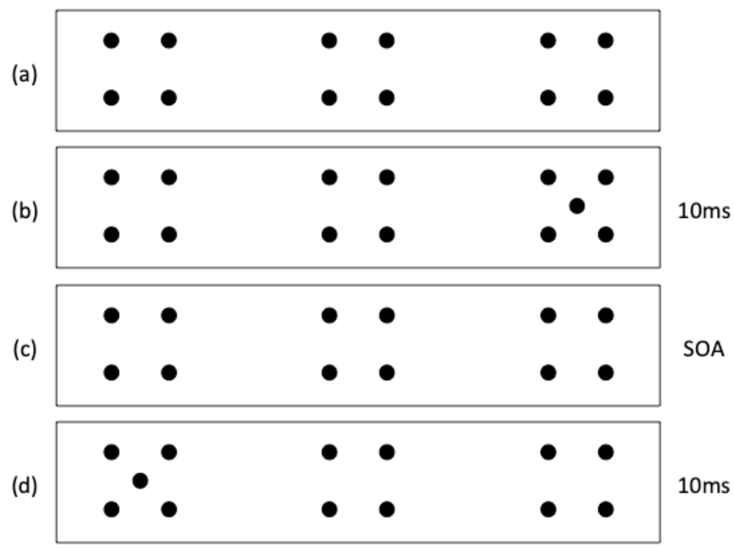


Figure 4. Diagram of the series of events from Stelmach and Herdman (1991). (a) Attention could be directed to the left, right, or center square. (b) A target dot appeared in the left or right square for 10ms. (c) A variable SOA (i.e., 0 – 70m) in which no targets were presented. (d) The second target appears on the opposite side for 10ms.

When participants were asked to attend to the middle box, the point of greatest temporal uncertainty (i.e., the point at which participants were unable to tell which target appeared first) was around an SOA of 0ms meaning that the peripheral targets appeared at roughly the same time and were perceived as such. When participants were asked to attend to either the left or the right square, the stimulus occurring at the *unattended* location had to *precede* the attended side by 40ms in order for the point of greatest temporal uncertainty to be reached. That is, the point of subjective simultaneity (PSS) was observed when the unattended side was presented 40ms prior to the target on the attended side. These findings suggest that directed attention results in facilitated processing of information for items occurring at the attended location. Stelmach and Herdman (1991) argued that this leads to shorter transmission latencies for information appearing at the attended location compared to information appearing at the unattended location. Thus, in order for the two stimuli to be perceived as occurring simultaneously, the *unattended* side must appear *before* the attended side by about 40ms.

Further evidence for the facilitatory role of attention in information processing can be seen in tasks involving perceptual learning (Gibson, 1969), wherein the ability to perceive and respond to stimuli is enhanced through experience. Ahissar and Hochstein (1993) asked participants to view arrays of slanted lines while performing one of two different tasks: global detection or local detection. In the global detection task, participants were asked if the array itself was oriented in a vertical or horizontal configuration while ignoring how the stimuli contained within the array were arranged. In the local detection task, participants had to determine if one of the lines within the array was oriented in a different direction from the rest of the lines in a standard “pop-out” search task (Treisman & Gelade, 1980, Figure 5). Here, attention was directed to the specific stimuli within the array, while the overall orientation was ignored. Participants were trained on one task (i.e., global or local detection) until they reached a performance threshold (i.e. performance improvement as a function of session number reached asymptote). Afterward, they were tested on either the same task using different stimuli (i.e., stimulus specificity) or the other task using the same stimuli (i.e., task specificity).

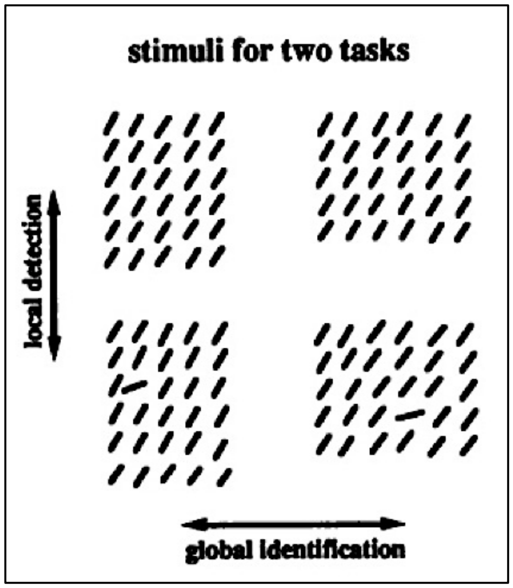


Figure 5. Example of array configurations in Ahissar and Hochstein (1993). The local detection task is depicted on the y-axis – participants were asked to locate the "odd-ball" orientation within the array of oblique lines. The x-axis depicts the global identification task – participants were asked to determine if the array itself is oriented vertically (right side) or horizontally (left side).

In the stimulus specificity condition, performance on both detection tasks (global or local detection) declined when the stimuli were changed compared to when the stimuli remained the same. However, Ahissar and Hochstein (1993) also found that perceptual learning contains a task-specific element. Participants trained on local detection had comparatively worse performance on global detection and vice versa when the stimuli remained the same. This finding suggests that, in addition to low-level retinotopic training with specific stimuli typically seen in perceptual learning tasks, learning proceeds according to higher-level, extraretinal control mechanisms that may be linked to attention. Therefore, learning likely begins with processing information in a low-level, bottom-up, manner that is likely influenced by descending top-down attentional signals that mediate task performance according to task demands.

As the research outlined in this section demonstrates, directing attention to a particular spatial location or stimulus element can lead to facilitated processing for the attended information. This can enhance processing speed, as in the TOJ task, or learning of a particular stimulus feature, as in the perceptual learning task. While these findings demonstrate the facilitatory role of attentional processes for items that we explicitly direct our attention towards, facilitation can also occur for items that are explicitly ignored as well, demonstrating that information processing may be aided by mechanisms beyond attentional capacity such as temporal presentation (or alignment) with an item that *is* being attended.

3.3.1. Facilitation for Explicitly Ignored Information

In our complex environment, the items or locations we attend to do not occur in isolation. That is, we must *select* items of relevance from an array of potential candidates. Within this framework, one question that remains is, what happens to the information that is ignored? Arguably, from an information processing perspective, *something* must be done with that information. Indeed, even if we focus our attention toward a particular stimulus – a road sign, for example – we may also become aware of the information surrounding that stimulus (buildings, people, or other landmarks), even if our attention was not explicitly focused on those items. Thus, a growing body of research has aimed at 1) determining the extent to which unattended information is processed, 2) identifying the factors that may increase the likelihood that unattended information is processed, and 3) how these events might impact behavior. Typically, the degree to which unattended, task-irrelevant, stimuli are processed has been investigated using paradigms employed in the exploration of visual perceptual learning and inattentional blindness (Dewald & Sinnett, 2011a, 2011b; Dewald et al., 2011, 2013; Rees et al., 1999; Seitz & Watanabe, 2003, 2005; Sinnett et al., 2006; Tsushima et al., 2006; Tsushima et al., 2008; Walker et al., 2014; Walker et al., 2017; Watanabe et al., 2001). In these paradigms, participants are required to first engage in a primary task wherein they must monitor specific stimuli for the appearance of a target while ignoring simultaneously presented task-irrelevant information. The extent to which the ignored items were processed, and subsequently remembered, is then assessed via some variation of a surprise recognition or identification task.

Findings from these studies have demonstrated that the temporal presentation – or more precisely, the temporal *alignment* – of the ignored items plays a crucial role in facilitating or inhibiting the extent to which these items are processed[2]. Collectively, this body of work demonstrates that when task-irrelevant items are *frequently* presented with targets (i.e., *target-aligned* or TA) in the primary task, then recognition for these items during the surprise recognition test is facilitated compared to the ignored items that were not presented with targets (i.e., *non-aligned* or NA) during the primary task but were presented with equal frequency. Watanabe et al. (2001) provided a seminal account of this phenomenon by asking participants to engage in an attention-demanding, target identification, primary task while ignoring a simultaneously presented task-irrelevant dynamic random dot (DRD) display in the background. In this experiment participants were asked to monitor a RSVP stream of black letters for a target grey letter, presented at fixation, while the DRD display was presented in an annulus around the fixation point (Figure 6).

Critically, within the DRD display, a small subset (5%, subthreshold to explicit awareness) of the otherwise randomly moving dots moved coherently in the same direction while participants engaged in the primary task. After completing the primary task, participants were given a surprise motion coherence identification task wherein they were asked to determine the direction of coherently moving dots in a similar DRD display with a subset motion coherence of 10%, which importantly was a coherence that afforded suprathreshold identification. Critically, participants were significantly better at detecting motion coherence during the surprise identification task when the dots moved in same direction as the subthreshold (5%) motion coherence they were exposed to during the primary task, compared to other motion directions they had not been exposed to.

---

[2] While this body of literature outlines the temporal conditions that may lead to inhibited or facilitated processing of task-irrelevant information (Dewald & Sinnett, 2011a; 2011b; Dewald et al., 2011, 2013; Rees et al., 1999; Seitz & Watanabe, 2003, 2005; Sinnett et al., 2006; Tsushima et al., 2006; Tsushima et al., 2008; Watanabe et al., 2001) this dissertation will only focus on conditions of facilitated processing.

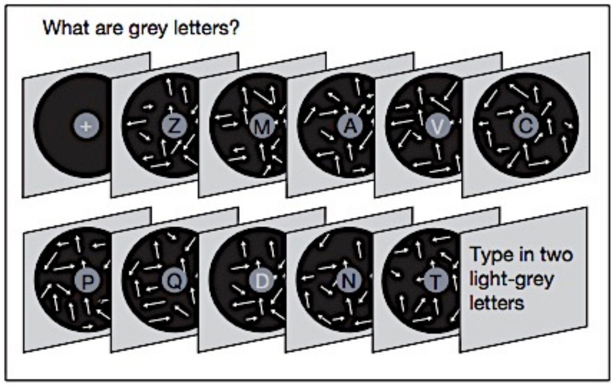


Figure 6. Example of RSVP task and DRD presented during the primary task of Watanabe et al. (2001). Participants monitored the letter stream for target grey letters (the 'V' and the 'D') while ignoring the DRD display presented around the fixation point.

Seitz and Watanabe (2003) extended this work by systematically varying the direction of subthreshold motion coherence appearing in the background DRD display such that participants were exposed to four different directions during the primary task. The critical manipulation in this experiment was that each time a target appeared in the RSVP stream, participants were always exposed to the same motion direction and when there was no target present, participant were exposed to a random presentation of the remaining three motion directions – though all participants were exposed to each of the four motion directions an equal number of times. Therefore, the motion direction appearing with the task-relevant targets was now target-aligned (TA) whereas the remaining three motion directions were non-aligned (NA). During the surprise motion coherence identification test, participants were significantly better at detecting the TA motion direction compared to the remaining three NA motion directions.

Dewald and colleagues (2012) extended this work with more complex and salient stimuli (i.e., written words and pictorial images), by utilizing a procedure employed by Rees et al. (1999; see also Sinnett et al., 2006) in which participants viewed a RSVP stream of pictures with words superimposed on top. Participants were required to focus their attention on the pictures (while ignoring the superimposed words) and identify an immediate picture repetition (i.e., the task-relevant target) in the RSVP stream. In order to align with stimulus presentations in Seitz and Watanabe (2003), only one unchanging word was paired with the picture repetition targets (i.e., TA) while an additional seven words were randomly presented over the non-repeating pictures (i.e., NA) – again all words were presented an equal number of times. After the primary task, participants were given a surprise recognition test in which they were asked to distinguish between words presented during the experiment and never before seen foil words in a yes-no recognition task (Jang, Wixted, & Huber, 2009).

Interestingly, recognition rates for both TA and NA words were significantly *above* chance. However, TA words were recognized at rates significantly higher than NA words. These findings suggest that the temporal alignment of unattended, task-irrelevant stimuli with task targets may play a crucial role in modulating the extent to which information that is actively ignored may be facilitated during information processing. This observed facilitation for TA items may arise because these unattended items are temporally and spatially aligned with attended task targets. The guided search model (Wolfe, 2001; Wolfe & Gray, 2007) supports this account given that when a target is 1) identified and 2) responded to, it is assumed that it has passed through any and all potential filters in the system allowing for the necessary identification and response selection processes to take place. As seen in the AB paradigms (Shapiro, 1994) unattended items may be processed in parallel with the attended information. Thus, despite attention not being directed explicitly *toward* those items, they also receive additional processing because they are presented with information that is attended and, more importantly, acted upon.

While it has been established that one potential mechanism involved in the facilitated processing of explicitly ignored, task-irrelevant, information may be temporal alignment, additional factors may also play a role in this process. For example, multisensory integration (i.e., the process of synthesizing information arriving at our divergent senses) has been shown to alter and sometimes enhance information processing under certain conditions for both attended and unattended information (Stein, Stanford, Wallace, & Jiang, 2004; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011). Thus, it stands to reason that the sensory modalities involved in stimuli presentation may play a role in facilitating processing for unattended items. The next chapter will explore some of the ways in which our perceptual system processes multisensory information and how this may be modulated by attentional processes.

4. Multisensory Integration

Our world is a complex multisensory environment and our perceptual system allows us to take in divergent streams of sensory information and combine them into a single unitary percept through multisensory integration (MSI). This intricate process of organizing and integrating various sensory streams helps us to perceive stimuli by reducing noise within our perceptual system. Noise reduction is accomplished through binding information from different senses together (Stein et al., 2004) thereby allowing us to interact, adapt, and navigate through our surroundings. Thus, MSI makes it easier to select potentially relevant pieces of information from background noise and isolate temporally successive events from one another. However, this process can be influenced by a number of various factors, which can, ultimately, alter our resulting perceptual experience. This section will first examine the ways in which incoming information from one sensory modality can influence our perception of information arriving via a different sensory modality; next, we will explore some of the ways in which attentional processes may interact with MSI to, again, alter our perceptions of the world around us.

4.1. Interactions Between Visual and Auditory Senses

While it is not contentious to assert that we have many sensory modalities, it should be noted that the majority of relevant research, so far, has focused primarily on visual and auditory processing, and only relatively recently, the resulting interactions between these two senses. As such, the topics introduced here will focus primarily on these two sensory modalities. Auditory and visual information frequently co-occur in the environment and their concurrence may hold relevance (e.g., they may originate from a single source) or they may not. When two, or more, sensory modalities provide information about the same property of the external world the processes of combining those streams of information is called *convergent MSI* (Driver & Spence, 2000), which is perhaps the most commonly thought of form of MSI. However, it is important to note that MSI does not just occur for meaningful or task-relevant information. Indeed, irrelevant information presented in one modality (e.g., auditory) can influence perceptions of, and responses to, information presented in a different modality (e.g., visual) even when the former provides no meaningful information about the latter. In this case, the properties of each sensory modality are considered orthogonal to one another (Driver & Spence, 2000), yet they may still be combined by our perceptual system and information from one sensory modality may influence our perception of the other.

One of the most well-known examples of how MSI can lead to alterations in our perceptual experience between the visual and auditory modalities can be seen in the McGurk effect (McGurk & MacDonald, 1976). This is an auditory perceptual illusion that arises from convergent MSI between an auditory signal and a concomitant, incongruent, visual signal – meaning that modifying the visual signal influences perception of the auditory signal. The illusion arises when an individual observes a speaker generating a simple monosyllabic sound like /ba/. Producing this particular sound requires specific movement of the external oral articulators in which the upper and lower lips are pressed together and turned slightly inward at the initiation of the sound followed by an opening of the mouth into a slight "O" shape. However, when this auditory signal is paired with an incongruent visual stimulus, the perception of the sound is altered. For example, the lip movements associated with the monosyllabic sound /ga/ begin with both lips parted, usually revealing several teeth, and the mouth slightly open. When the auditory signal /ba/ is paired with the lip movements for /ga/, the resulting auditory perception is /da/ despite the physical properties of the auditory signal remaining constant. This type of effect has also been noted in non-linguistic stimuli such as viewing one musical instrument while listening to the sounds produced by a different instrument (Saldaña & Rosenblum, 1993).

Another interesting influence of visual information on auditory perception can be seen in the ventriloquism effect (Bertelson, 1999; Thurlow & Jack, 1973). In this case, an auditory signal can be "pulled" to the spatial location of a visual event that is not generating the sound – thereby affecting sound localization. This is a common perceptual phenomenon experienced by moviegoers or individuals with an at-home surround sound entertainment system. In this case, viewers watch the actors on the screen while the sounds are generated from loud speakers positioned around the room, rather than directly ahead (as the screen is). Despite the sound originating from a location that is different from that of the visual signal, the resulting perception is that the auditory words and sound effects are emanating from the screen.

While it is well established that visual input can alter perception of an auditory signal, it should be noted that adding an auditory signal to a visual stimulus can alter how we perceive that visual information. Indeed, the ventriloquist effect can be bi-directional. In some cases, an auditory signal can be pulled toward the spatial location of a visual event (as described above) while under different circumstances an auditory signal can pull a visual event temporally. In this case, the perception of *when* a visual event occurred can be altered by the application of an auditory signal (Morein-Zamir, Soto-Faraco, & Kingstone, 2003). Another example of auditory influences on visual stimuli can be seen in the illusory flash effect (Shams, Kamitani, & Shimojo, 2000; see also Parker & Robinson, 2018). Here, a single visual flash is often perceived as multiple different flashes when a series of short auditory beeps are presented within 100ms of the flash onset.

A more recent example of how an auditory signal can alter perception of a visual event is called the pip and pop effect (Van der Burg et al., 2008; Van der Burg et al., 2011). Here, participants completed a complex visual search task in which the target (a vertical or horizontal line) and distractors (oblique lines) randomly change color from red to green. In one condition, participants simply completed the visual search task with no additional information. In the multisensory condition, a non-spatial auditory cue (i.e., a "pip") was played that coincided with the color change of the target only. When the color change of the target was temporally paired with the auditory pip, performance on the visual search task was greatly improved. Astonishingly, search performance switched from those typical of a conjunctive search task, in which the time to detect the target increases as a linear function of set size, to a singleton "pop-out" search, which returns a flat, or nearly flat, search function as set size increases (Treisman & Gelade, 1980). Van der Burg et al. (2008) suggested that the synchronous presentation of the auditory pip with the visual target added a measure of temporal resolution that the visual modality lacks. Therefore, the added auditory cue boosted the saliency of the visual target causing a salient emergent feature, resulting in the impression of the target "popping-out" of the display.

This sample of literature has explored some of the ways in which information presented to the auditory modality may alter perceptions of visual information and vice versa. However, there are additional factors that influence not only the extent to which multisensory information might be combined, but also the way in which it is combined. For example, how we direct our attention to a particular object or spatial location can influence the way in which our brains combine sensory information. The next section will examine some interactions between attention and MSI.

4.2. Interactions Between Multisensory Integration and Attention

Chapter three outlines some of the potential roles of attentional allocation when processing information in general, such as gating the flow of incoming information, inhibiting processing for task-irrelevant stimuli, and facilitating information processing for both attended and ignored items. However, attention also impacts the way in which we combine sensory information. In some cases, attention can spread across modalities such that attending to information presented in one sensory modality can increase responding to information presented in a task-irrelevant modality (Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Donohue, Roberts, Grent, & Woldorff, 2011). In other circumstances, cueing attention to a particular location in one sense can lead to a complementary shift in spatial attention of a separate sense (Driver & Spence, 1998; Spence & Driver, 1996; 1997; Störmer, McDonald, & Hillyard, 2009). Finally, attention can also affect the speed at which information arriving from a particular sensory modality is processed (Spence et al., 2001; Titchner, 1908). These forms of multisensory (AKA cross-modal) interactions will be explored in the following sections.

4.2.1. Cross-Modal Spreading

As we saw in chapter three, attending to a particular location can lead to facilitated RTs, increased stimulus discrimination, and enhanced perceptual accuracy (Posner, 1980). When it comes to MSI, attention researchers have also attempted to determine the extent to which attending to one sensory dimension (e.g., physical shape) of a multisensory event enhances processing for co-occurring, yet unattended, sensory dimensions (e.g., sound). Busse and colleagues (2005) investigated this type of cross-modal spreading in a study using event related potentials (ERPs) and functional magnetic resonance imaging (fMRI).

Busse et al. (2005) asked participants to focus on a central fixation point and covertly attend to either the right or left side of a computer screen while a RSVP stream of checkerboard displays appeared at either location (i.e., right or left). Participants were instructed to ignore all stimuli presented to the unattended side while they monitored the attended side for an infrequently occurring target checkerboard containing two dots (Figure 7). On 50% of the trials, the visual stimuli were presented with a task-irrelevant auditory signal (i.e., a pip) that was generated from a central location behind the monitor.

The authors found that when the attended visual target was accompanied by the irrelevant auditory signal, participants were significantly more likely to detect the target. This finding corroborated ERP results revealing enhanced neural responding to the tone when it was combined with attended visual stimuli, compared to being paired with unattended stimuli or presented alone. The ERP reflection of the enhanced responding was a sustained, frontally distributed, brain wave that emerged late in stimulus processing (~220ms after stimulus presentation). The distribution of the brain wave was consistent with major contributions from the auditory cortex and the time course of the enhancement suggests an attention-related component (Busse et al., 2006). fMRI data further corroborated these findings by revealing enhanced activity in the auditory cortex in response to tones paired with attended visual stimuli, compared to tones paired with unattended stimuli or presented alone.

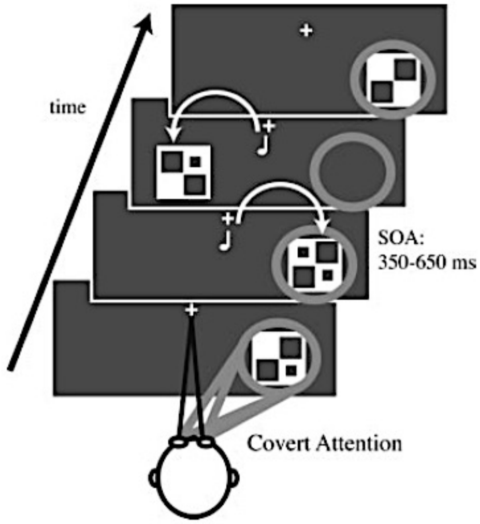


Figure 7. Example of stimuli and task from Busse et al. (2006). Participants covertly monitored one side of the screen (indicated by grey circles) for a target checkerboard containing two black dots, while ignoring all stimuli presented to the unattended side. Half of the visual stimuli were accompanied with a task-irrelevant auditory signal.

The results from Busse et al. (2005) suggest that simultaneously presenting attended visual stimuli with unattended auditory stimuli leads to the spread of attention across sensory modalities and space to encompass the task-irrelevant auditory signals. The sensory signals from the unattended auditory modality, positioned in a separate spatial location, are grouped with the synchronously presented and attended visual sensory input into a single multisensory object. As a result, the unattended modality is pulled into the attentional spotlight where processing of this information is enhanced. This spreading activation appears to enhance late neural activity in response to the additional sensory dimension even if that information is irrelevant to task completion. Behavioral results also suggest that this process boosts the signal of the attended stimuli thereby facilitating detection and responses along the task-relevant domain.

4.2.2. Cross-Modal Cueing

In addition to spreading attention across sensory modalities, the allocation of attention can be cued across the senses. In this case, providing an endogenous or exogenous cue to one modality may facilitate attentional orienting in a separate sensory modality. Spence and Driver (1996, 1997) explored this in a series of seven experiments wherein attention in either the visual or auditory modality was cross-modally cued under exogenous and endogenous conditions. While a comprehensive review of this literature is beyond the scope of this dissertation, experiments 2 and 3 from Spence and Driver (1996) can help elucidate the concept of cross-modal cueing.

Spence and Driver (1996) first examined cross-modal cueing in endogenous (i.e., voluntarily controlled) covert orienting. Their experiments utilized a multisensory variation of the classic cueing paradigm (Posner, 1980). Participants fixated on a dot of light positioned directly in front of them. Two loudspeakers – one on top of the other – were positioned to the left and right side of the participant; each speaker had a LED placed in its center (see Figure 8). Thus, the loudspeakers generated auditory targets and the corresponding LEDs generated visual targets from the same spatial location. Most trials began with an informative central visual arrowhead cue indicating the most likely side a target would appear (left or right). Participants were told to keep fixation on the central dot while covertly orienting (i.e., without moving their eyes) their auditory (Experiment 2) or visual (Experiment 3) attention according to the arrowheads that appeared at the start of each trial. In both experiments, participants were asked to respond to auditory *and* visual targets. Regardless of cue type (valid, invalid, or neutral) or target modality (visual or auditory) participants had to make an elevation judgment (up vs. down) for the stimuli occurring on the left or right.

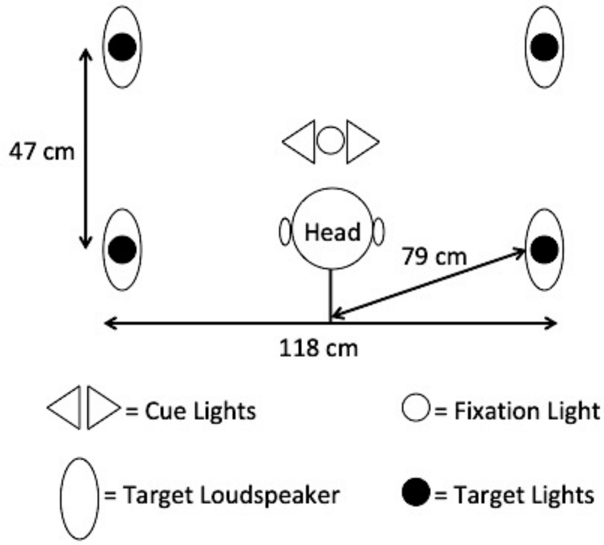


Figure 8. Diagram of the stimuli in Spence and Driver (1996). Loud speakers were positioned, one on top of the other, to the left and right of the participant. Each loud speaker had a LED placed in the center. Participants maintained eye gaze on the central fixation light for the duration of the experimental session. Cue lights indicated the most likely side a target (auditory or visual) could appear. Cues could be valid, invalid, or neutral.

Results from Experiment 2 (auditory attention with visual cues) revealed that participants were able to shift endogenous auditory attention to a spatial location indicated by a visual cue leading to faster auditory target detection at the cued side. Visual attention also tended to shift with auditory attention toward the cued side leading to faster RTs to visual targets despite the fact that they were far less likely to appear. Spence and Driver (1996) took this finding as preliminary evidence that visual cues facilitate spatial shifts in auditory endogenous attention and that orienting auditory attention to a particular location tends to be accompanied by complementary shifts in visual attention in the same direction. Results from Experiment 3 (visual attention with visual cues) dovetailed with the results of Experiment 2. When participants directed their endogenous covert visual attention to a spatial location, auditory spatial attention tended to follow. Combined, the findings from the two experiments suggest that, while cueing effects tend to be largest for the attended sensory modality, endogenous visual attention tends to follow endogenous auditory attention and vice versa. Based on findings from all seven experiments, the authors concluded that visual and auditory attention neither operate in a purely supramodal (i.e., combined sensory) or a purely independent (i.e., modality specific) fashion. Rather, it is likely that there exist separable systems for audition and vision that work cooperatively in spatial synergy. Covert attention between the two modalities can be more efficiently applied to the same location rather than to two opposing locations. This means that, unless circumstances strongly dictate otherwise, attention between the two sensory modalities will be directed to the same spatial location, which enhances detection of visual and auditory stimuli.

4.2.3. Cross-Modal Prior Entry

In addition to the spreading of attention from one sensory modality to another leading to spatial orienting effects across the senses, attention can also influence the speed at which information arriving through our different sensory modalities is processed. Spence and colleagues (2001) explored the extent to which directing attention to a particular sensory modality could influence *prior entry* (Titchner, 1908) for that sensory information. The authors hypothesized that directing attention to a particular sensory modality could facilitate the speed at which that information was processed compared to information simultaneously presented to an unattended sensory modality.

In their experiments, Spence et al. (2001) presented participants with a modified TOJ task using visual and tactile stimuli. Subjects were positioned in front of a central fixation LED flanked by two additional LEDs, one to the left and one to the right. Participants were asked to place each of their index fingers on a stationary vibrator positioned on their left and right sides (i.e., in corresponding locations to the peripheral LEDs). Participants were presented with both stimulus types (visual and tactile) on each trial in various combinations: visual and tactile on the left, visual and tactile on the right, visual on the left and tactile on the right, and visual on the right and tactile on the left. In order to examine prior entry for the two different sensory modalities, participants were asked to determine which modality was presented first (i.e., modality TOJ). This task was conducted while participants divided their attention across both sensory modalities (i.e., visual *and* tactile) or focused their attention on one modality only (i.e., visual *or* tactile).

When attention was divided across visual and tactile modalities, the visual stimulus had to be presented *before* the tactile stimulus by 30ms in order for participants to reach the point of subjective simultaneity (PSS, i.e., the point at which they perceived the two stimuli as co-occurring in time). Participants also tended to be more accurate when the stimuli were presented from separate spatial locations. These findings suggest that when attention is divided between visual and tactile inputs, processing for visual information may lag slightly behind tactile information. The fact that the PSS was less accurate when stimuli were presented on the same side, compared to opposite sides, suggests a sensory integration component. When tactile and visual stimuli occurred in the same spatial location, the signals may become bound together making temporal order discriminations more difficult (Spence et al., 2001).

When participants focused their attention to one sensory modality, the authors noted significant prior entry effects. Directing attention to the tactile modality increased the speed at which this type of information was processed. As a result, visual stimuli had to lead by a greater interval (155ms), compared to the divided attention condition, in order for the PSS to be achieved. Conversely, when attention was directed to the visual modality, the PSS shifted in favor of visual processing. Here, visual stimuli had to lead tactile by 22ms, which was a significant reduction from the condition in which attention was directed to tactile input as well as when attention was divided across both sensory modalities (i.e., 30ms). According to Spence et al. (2001) these results clearly indicate that attending to one sensory modality speeds the perception of stimuli occurring in that modality compared to conditions in which attention is directed to a separate modality. The fact that attending to the tactile modality resulted in stronger prior entry effects was attributed to the notion that people have a general bias toward the visual modality under conditions of divided attention, thus directing attention to that modality has less of an impact on processing speed.

Chapters three and four outlined the roles of attention in the information processing system, the potential consequences of directing attention either toward or away from a particular stimulus, and the influences of attention on MSI. Specifically, attention can help to control the flow of incoming information into a limited capacity system by directing the selection of relevant items for further processing (Shapiro, 1994; Treisman & Gelade, 1980; Wolfe & Gray, 2007). In doing so, information from an unattended source is likely to be inhibited resulting in diminished re-orienting to a previously attended location and slower reaction times (Posner & Cohen, 1984), as well as lower rates of stimulus perception and poorer memory recall (Mack & Rock, 1998a, 1998b; Neisser & Becklen, 1975; Simons & Charbis, 1999). However, directed attention can also facilitate information processing for both attended (Ahissar & Hochstein, 1993; Stelmach & Herdman, 1991) and ignored information (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe et al., 2001) thereby increasing the likelihood that information will be perceived and recognized later. Furthermore, attentional processes may interact with sensory modality such that attending to information in one sensory modality can enhance processing of co-occurring, yet task-irrelevant, information from a separate modality in the form of attention driven spreading activation (Busse et al., 2005). Attentional orienting within our sensory systems (e.g., vision, touch, and audition) tends to operate in a synergistic manner such that one can facilitate the other (Spence & Driver, 1996). Finally, directed attention can enhance the speed at which information from an attended modality is processed (Spence et al., 2001).

Collectively, this body of work provides compelling evidence for the role of attention in an information processing system, with specific focus on attentional allocation operating to facilitate the processing of both attended and ignored information and subsequent interactions with MSI. However, the extent to which information may be processed is likely influenced by a number of factors. In addition to attentional allocation and sensory modality, it is plausible that the stimulus type *itself* may play a central role in information processing and that this factor may interact with attentional processes and sensory modality. Chapter five explores two important types of information we encounter, some of the ways in which information processing for these two stimulus types might be similar or different, and how processing might be affected by attentional allocation.

5. Attention and Processing Lexical and Non-Lexical Information

Chapter four outlined the multisensory nature of our complex environment, however it should be noted that in addition to receiving information through our various senses, the *type* of information each sensory modality receives can also vary. As such, the way in which our cognitive system handles the different forms of information we encounter may be different and this may be influenced by sensory modality. Arguably, two important types of information we encounter on a daily basis, and process through our visual and auditory senses, include semantic lexical (i.e., words both written and auditory) and semantic non-lexical information (i.e., visual imagery and sounds). Historically, research exploring potential differences in processing lexical vs. non-lexical information has focused on conditions in which these types of stimuli are actively attended. However, the extent to which observed differences in information processing under attended conditions might generalize to conditions of inattention remains underexplored. This section will sample some of the research that has attempted to understand how our cognitive system processes lexical and non-lexical information under conditions of directed attention, as well as available research under conditions in which one, or both, types of stimuli might be actively ignored.

5.1. Attended Lexical and Non-Lexical Information

5.1.1. Visual Modality

An existent body of research suggests that when we focus our attention toward a specific stimulus, the way in which that information is processed may be influenced by the kind of information that stimulus represents. Specifically, several studies have observed behavioral differences in the processing of visual lexical information, such as words, and visual non-lexical semantic information, such as pictures, (Amit et al., 2009; Carr, McCauley, Sperber, & Parmelee, 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980). Furthermore, these differences are observed despite distinct overlaps in the neurological pathways that are believed to be responsible for processing each stimulus type (Bright, Moss, & Tyler, 2004; Vandenberghe et al., 1996). Amit et al. (2009) suggested that differences in processing between lexical and non-lexical information arise because words serve as arbitrary generic labels for categories of items and that because of this lack of specificity words maintain a higher level of conceptual abstraction. For example, the word *car* describes a class of items that can, and do, vary drastically from one another (i.e., a Toyota Prius and a Lamborghini Veneno Roadster are both "cars" yet they bear striking differences). Therefore, comprehension of the meaning of the word *car* is a semantic construct that may vary drastically from person to person and instance to instance. Conversely, semantic images, such as pictures, are explicit, highly specific, self-contained items in and of themselves. This high level of concrete specificity may result in relatively low levels of conceptual abstraction. Returning to our example of cars, while the word may be relatively abstract, an image of a Toyota Prius or a Lamborghini Veneno Roadster provides a literal example of a particular type of car with a distinct shape and size, thereby removing the need to "decode" or interpret the information. Therefore, it is likely that processing pictorial semantic information may be facilitated, compared to written words, due to a lower level of conceptual abstraction.

The relationship between stimulus type (i.e., words vs. pictures) and information processing has been examined in several studies to date (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980). Specifically, this body of research has examined the conditions that lead to facilitated processing for each stimulus type (pictures or words). For example, Carr et al. (1982) briefly presented participants with either a semantic picture – depicting a common object (e.g., a spoon) – or a word at fixation and asked them to make one of two judgments: to either name the item or categorize it. Additionally, the target item (i.e., picture or word) could be preceded by a prime word or picture that was either related or unrelated to the target item. In general, within stimulus primes (e.g., a picture prime presented before a picture target) led to faster RTs compared to cross-stimulus primes (e.g., a word prime presented before a picture target) and responses to picture targets were more likely to be facilitated by a prime (regardless of prime type) compared to responses to word targets. Furthermore, for all conditions, when the target items (words or pictures) must be categorized, participants were faster for pictures compared to words. On the other hand, when the target items must be named, participants were faster to respond to written words compared to pictures (Carr et al., 1982; see also, Hogaboam & Pellegrino, 1978; Smith & Magee, 1980).

The observed differences in processing between pictures and words may be explained by how our cognitive systems handle each type of stimulus. Categorization for pictures is facilitated compared to naming these items because pictures maintain more direct access to semantic information. In order to categorize an item it must first be recognized and object recognition is typically considered to involve the retrieval of visual and abstract semantic information, which occurs prior to the retrieval of lexical or name representations (Grill-Spector & Kanwisher, 2005; Humphreys & Forde, 2001). Words may take longer to categorize because their higher level of conceptual abstraction requires that they must first be decoded before they can be semantically evaluated (i.e., they must be read first and then associated with a meaning). On the other hand, naming may be faster for words because accessing semantic information is not necessary for reading a word, while it can be argued that this must still occur in order to name a picture (Carr et al., 1982; Stroop, 1935). In sum, Carr and colleagues (1982) suggest that the less effort our cognitive system must put in to transforming incoming information from one code to another, the faster that transfer process will be, and this process may be influenced by both the type of stimulus and demands of the task.

5.1.2. Auditory Modality

In addition to the visual modality, semantic lexical and non-lexical information can be processed via our auditory system. It should be noted that a comprehensive comparison between vision and audition in general and, specifically, with regard to the various stimulus types (lexical vs. non-lexical) each modality can processes is beyond the scope of this dissertation. Furthermore, while some researchers have explicitly examined how lexical and non-lexical information might be processed via the visual modality there are far fewer investigations examining the precise differences in how these two forms of information might be processed via the auditory modality. Broadly speaking, auditory word recognition is thought to progress along two main levels of representation: lexical and sub-lexical (McClelland & Elman, 1986; Vitevitch & Luce, 1998, 2016). At the lexical level of processing, speech sounds produce corresponding lexical/semantic representations that compete for activation (Colombo, 1986; Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994). While the precise nature of sub-lexical processing is less agreed upon, there exists a general consensus that the individual components of an auditory word may also correspond to independent representational entities that are activated during speech perception (McClelland & Elman, 1986; Vitevitch & Luce, 1998, 2016).

Contributions from lexical and sub-lexical components may influence the extent to which the sound signal is processed. For example, Vitevitch and Luce (1998) examined the effects of probabilistic phonotactics (i.e., word segment and syllable frequency) and similarity neighborhood-density (i.e., the amount of competition from corresponding lexical/semantic representations that may be activated by a word or phrase) on linguistic processing. Probabilistic phonotactics is believed to represent sub-lexical properties while neighborhood-density is believed to represent lexical properties associated with the speech sounds. In their experiment, participants were placed in front of a boom microphone and fitted with a pair of headphones. At the start of each trial, participants were presented with either a word (i.e., lexical processing) belonging to high-density (i.e., words that sound similar to many other words, such as “rate”) or low-density (words that sound similar to only a few other words, such as “bulb”) representational systems (Luce & Pisoni, 1998), or non-words consisting of speech sounds (i.e., sub-lexical processing) of either high probability (i.e., sounds that occur frequently in everyday speech such as /*kik*/) or low probability (i.e., sound that are relatively rare in everyday speech such as /*guil*/) via the headphones and asked to repeat the word or speech sound as quickly and accurately as possible.

The authors found that probabilistic phonotactics and neighborhood-density have differential effects on processing linguistic information, which is further influenced by lexical status. Specifically, lexical processing is inhibited (i.e. participants were slower to respond) for words occurring in high-density neighborhoods compared to words occurring in low-density neighborhoods despite high-density words containing sub-lexical components with high probability phonotactics. On the other hand, sub-lexical processing was facilitated (i.e., participants were faster to respond) for non-words with high-probability phonotactics compared to low-probability phonotactics, despite the fact that these speech sounds are typically associated with words occurring in high-density neighborhoods.

Vitevitch and Luce (1998) interpreted these orthogonal results to suggest that words are associated with multiple (and potentially competitive) semantic nodes stored in memory while non-words are not. Therefore, high-density words activate a larger number of representational competitors with divergent semantic links thereby inhibiting processing for those items to a greater extent compared to non-words with similar properties. For example, hearing the word "rate" may also activate representational competitors such as "late," "date", or "ate" which have similar phonetics (i.e., they sound similar) but divergent semantic connections within the representational system, which all compete for activation. Furthermore, a single high-density word such as "rate" may be tied to divergent semantic links. For example, the word "rate" can be used to indicate speed (e.g., heart *rate*), quality (e.g., to *rate* a product), or quantity (e.g., interest *rate*), which also compete for activation. On the other hand, because non-words do not activate representational nodes, higher activation associated with the sub-lexical units themselves (e.g., the frequency of the speech sound) lead to facilitated processing of the sound features when semantic information is limited or absent. Thus, common, sub-lexical, speech sounds such as /*kik*/ are more readily reproduced compared to less common speech sounds such as /*guil*/ (Vitevitch & Luce, 1998, 2016).

Thus, it is important to consider not only holistic differences in stimuli between the visual and auditory modalities but, also, the sub-components that comprise each stimulus type. Arguably, the different features of each stimulus type will lead to differences in how that information is processed by our cognitive system and, ultimately, how we respond to them. An additional factor of note includes our attentional allocation and how processing for each type of stimulus, in both the visual and auditory modality, may vary as a function of attention. To this point, we have only considered visual and auditory information to which attention has been directed. Another important question to consider is how might processing for information presented to both the visual and auditory modalities be affected when attention is actively diverted away from these stimulus types?

5.2. Ignored Lexical and Non-Lexical Information

5.2.1. Visual Modality

While some research has focused on potential differences in processing visual lexical and non-lexical stimuli when they are attended, it is not well understood how these types of stimuli may be processed differently when they are actively ignored. Indeed, most available research investigating the fate of ignored visual information using these stimulus types have focused exclusively on recognition performance of ignored words, using the pictures only as the attended stimuli in the primary task (Dewald et al., 2013; Rees et al., 1999; Ruz, Wolmetz, Tudela, & McCandliss, 2005; Ruz, Worden, Tudela, & McCandliss, 2005; Sinnett et al., 2006; Walker et al., 2014). Still other investigations have examined the effects of pairing irrelevant images with attended tasks (Swallow & Jiang, 2010, 2011) although the distractor images were explicitly attended along with the primary identification task. Therefore, these studies do not examine the extent to which actively ignored items may be processed.

Tipper and colleagues (Tipper, 1985; Tipper & Driver, 1988) focused on determining the extent to which ignored pictures and words might be processed. They presented participants with a combined “prime” stimulus containing either two pictures or two words (each one superimposed on top of the other). One of the combined items was always presented in green ink while the other was presented in red ink. Participants were asked to attend to the items presented in red while ignoring the items presented in green. After identifying the attended red item in the prime stimulus, participants were then presented with a "probe" trial that was also a combination of either two pictures or two words (combined stimuli were superimposed on top of each other, one in red and one in green, just as they were during the prime stage). When the probe appeared, participants were asked to categorize the attended red item as quickly and accurately as possible (explicit category labels were provided for participants at the start of the experiment). Next, they were asked to recall the category of the attended red item from the prime phase – this was done to ensure participants actually paid attention to the red items during the prime phase. Critically, the attended red probe could either 1) be in the same categorical domain as the ignored green portion of the prime (e.g., ignored picture/word prime followed by attended picture/word probe of the same category or vice versa), 2) share the same name as the ignored green portion of the prime (e.g., ignored picture/word prime followed by attended picture/word probe of the same name or vice versa), or 3) share no relation to the prime, (i.e., the control condition, prime and probe share no categorical relationship and do not have the same name, see Tipper & Driver, 1988 for full details of experimental design).

Results from their studies demonstrated that both ignored words and pictures are able to elicit negative priming (see Marcel, 1980; Neill, Lissner, & Beck, 1990; Tipper, 1985; Tipper & Driver, 1988; Tipper, MacQueen, & Brehaut, 1988) across stimulus types: meaning that ignored words can lead to slower response times for picture probes that pertain to the same category and vice versa. Specifically, participants were slower to respond to the attended probe (either picture or word) when it was preceded by a previously ignored prime (either word or picture) of the same category, compared to control conditions wherein ignored primes and attended probes shared no categorical relationship. This finding suggests that semantic information is indeed accessed for both pictures and words when these items are actively ignored. Thus, even under conditions in which attention is directed elsewhere, both stimulus types may be processed somewhat extensively, resulting in the formation of abstract representations that may influence behavior. This finding is further supported by later research conducted by de Zubicaray, McMahon, Eastburn, Pringle, and Lorenz (2006) demonstrating activation of the anterior temporal cortex, an area widely thought to be involved in semantic processing, in response to ignored primes in a similar negative priming paradigm. Though, it should be noted that this work involved the investigation of pictures only and did not address neural correlates associated with word processing in negative priming paradigms (see Deacon, Hewitt, Yang, & Nagata, 2000; Ruz et al., 2005, for examples of N400 processing for ignored words using event related potentials, ERPs). Still further evidence supports the notion that when pictures are actively ignored they may be identified beyond their physical features resulting in activation of phonological processing regions such as Wernicke's area and the posterior inferior frontal gyrus, implying that the *name* of the item may also be retrieved by linguistic areas (Bles & Jansma, 2008). While initial investigations reveal similarities in the extent to which both visual words and pictures may be processed when attention was focused elsewhere, much is left unknown. Still less in understood about how lexical and non-lexical information that is actively ignored might be processed by our auditory system.

5.2.2. Auditory Modality

Direct comparisons between unattended lexical and non-lexical information are difficult to find in the existent literature on information processing in the auditory modality. However, many studies have examined the extent to which unattended auditory words might be processed while engaging in an attention-demanding task. The most notable paradigm used to investigate the fate of ignored auditory lexical information includes variations of the dichotic listening task, also known as the cocktail party phenomenon, or inattentional deafness (Cherry, 1953; Conway, Cowan, & Bunting, 2001; Dalton, & Fraenkel, 2012; Giraudet, St-Louis, Scannella, & Causse, 2015; Haykin & Chen, 2005; Macdonald & Lavie, 2011; Murphy & Greene, 2015; Raveh & Lavie, 2015; Treisman, 1960).

In a dichotic listening task, participants are fitted with a pair of headphones and each ear is presented with a different sound stream. They are asked to attend to and shadow (i.e., repeat aloud) one sound stream and ignore the other. After this task is completed, participants are given a recognition test for the unattended ear. This recognition test is administered to evaluate the extent to which the ignored information may be processed and subsequently stored in long-term memory. Typically, recognition for the auditory information presented to the ignored ear is relatively poor, suggesting inattentional deafness, indicating that the unattended message was either inhibited (i.e., filtered out) at some point during information processing or subjected to rapid decay in the system resulting in a storage failure. Thus, participants are "deaf" to the auditory information due to a lack of attention. However, these investigations have demonstrated that, under certain circumstances, explicitly ignored lexical auditory information can be processed extensively, reaching a level of conscious awareness, resulting in explicit recognition of this information.

For example, the seminal study conducted by Cherry (1953) revealed that participants could detect certain featural (i.e., physical) changes in the unattended sound stream, such as when the speaker changed from male to female, or when the volume of the sound was increased or decreased, but failed to notice more complex semantic changes, such as when the message of the ignored ear was played backwards or changed languages. Thus, it appears that unattended auditory information may be subject to low level pattern recognition (such as pitch and amplitude), and arguably processed in parallel, while more complex semantic elements may be largely filtered out. However, some unattended auditory information, such as one's own name, may be weighted more heavily in our cognitive system, thereby bypassing the filtering restraints that are broadly applied to the auditory message. This suggests that some level of semantic analysis may be applied to the incoming information prior to being filtered but that only highly relevant or important information may be allowed to pass through to later processing stages.

A more recent investigation into processing unattended auditory information was conducted by Dalton and Frankel (2012), which aimed at evaluating the extent to which sudden onset, readily perceivable, sustained, lexical auditory information may be processed in a dynamic, three-dimensional, auditory scene that mimics a naturalistic setting. This study was conducted to compare instances of inattentional deafness with previous accounts of inattentional blindness, which mimic naturalistic environments (see Simons & Chablis, 1999). Listeners were presented with a binaural acoustic scene previously recorded in a dynamic three-dimensional space containing two sets of conversations centered on preparing for an upcoming party, one carried out between two women and one carried out between two men. During recording, both pairs of individuals (i.e., men and women) moved around the space as they conducted their conversation, thereby mimicking a real-life scenario in which multiple conversations are taking place and moving around in a three-dimensional space. Thus, when participants listened to the recording, both pairs of conversation were presented to each ear and both conversations "moved" around as the auditory recording was played. Listeners were instructed to attend to either the conversation of the women or the men. Half way through the recording, a third, unexpected, male voice was introduced that also moved through the auditory space while repeatedly saying the phrase, "I am a gorilla" for approximately 19 seconds. After listening to the auditory scene, participants were asked two questions (1) Did you hear anything unusual that didn't fit in with the scene? (2) Did you hear anyone other than the four people preparing for the party? If the subject answers 'yes' to either of these questions, they were further probed and asked to give more detail about what they heard.

Results from this experiment demonstrate clear inattentional deafness for the unexpected auditory information, but only under certain conditions. When participants were asked to attend to the conversation being carried out by the men in the auditory scene, 90% spontaneously reported hearing the gorilla upon being asked the first question. In contrast, reports of noticing the gorilla were much lower for participants instructed to listen to the conversation being carried out by women. Here, only 30% of listeners reported being aware of the gorilla, suggesting a significant level of inattentional deafness for the unexpected auditory message. These findings dovetail with results from classic dichotic listening tasks, suggesting that auditory lexical information may be largely filtered out when attention is directed elsewhere. However, unattended information may still be evaluated at a sub-lexical, featural, level leading to situations in which the ignored auditory signal may be facilitated and subsequently processed more extensively. Here, when participants were asked to attend to the men's conversation, the unattended auditory signal (i.e., "I am a gorilla") shared specific featural dimensions with the attended auditory signal (namely, pitch). This may have caused the unexpected irrelevant auditory information to be weighted more heavily in the system thereby allowing it to pass through the filter and receive additional processing (i.e., it was facilitated) and, ultimately, evaluated for semantic content. Interestingly, the extent to which this may occur may also be subject to specific spatial constraints, such as where the message appears in space relative to the attended and unattended message (Dalton & Frankel, 2012).

The research sampled thus far has only focused on the fate of unattended auditory lexical information, but has not addressed the question of what happens to auditorily presented, non-lexical, information that is not attended to. As with the visual modality, there are far fewer studies focused on understanding how unattended, non-lexical, auditory information might be processed. However, there appears to be growing interest regarding this kind of information. To this end, Koreimann, Gula, and Vitouch (2014) provided the first investigation of inattentional deafness in music perception. In their study, participants were presented with a brief musical clip from Richard Strauss' classical piece, *Thus Spake Zarathustra,* lasting approximately one minute and fifty seconds. During the last twenty seconds of the auditory clip an unexpected electric guitar solo was introduced to the sound stream. Half of the participants were instructed to count the number of tympani beats in the musical stream while the other half simply listened to the musical clip. After listening to the musical clip, participants were asked three questions of increasing specificity, (1) if they had noticed anything peculiar, (2) if they had noticed any unfitting instruments or sounds, and (3) if they had noticed the electric guitar. If participants answered ‘yes’ to any of the three questions, they were then probed for further information regarding what they heard, how it sounded, and when it occurred in the musical clip.

Of those participants who were instructed to count tympani beats, only 4% of participants explicitly noticed the additional electric guitar solo, while approximately 20% noticed something peculiar but could not explicitly identify what they heard. When participants were instructed to simply listen to the musical clip, the number of individuals who explicitly noticed the electric guitar solo increased to 52% of listeners while those who noticed something peculiar remained around 20%. This finding demonstrates that unattended, non-lexical, auditory information may be subject to similar processing constraints as lexical information. While some elements of the ignored sound signal may be evaluated at the featural level (i.e., pitch, amplitude, and timbre), thereby leading to partial or incidental recognition, the amount of information passed forward through the system remains limited. However, additional research must still be conducted before definitive conclusions may be drawn. The sixth chapter of this dissertation will outline the current project, which aimed at addressing some of the observed gaps in scientific literature.

6. Research Direction and Aims

Despite the rich and detailed investigations that have been conducted on processing unattended information thus far, there is a significant lack of research that has focused on explicitly comparing the extent to which different types of unattended information may be processed and how information processing might be affected by the interplay between stimulus type and sensory modality. Furthermore, the majority of available research has examined the rate of *inhibited* processing for task-irrelevant information, with comparatively few studies having focused on conditions that lead to *facilitated* processing. Developing a better understanding of how ignored information may be facilitated, and how this facilitation may be influenced by 1) the type of stimulus presented and 2) the modality of presentation, are important pieces to developing a comprehensive understanding of how the human mind evaluates, organizes, and prioritizes information in a dynamic and multisensory world.

The experiments for this dissertation addressed this gap in the scientific literature by explicitly and systematically comparing the impact of these dimensions (stimulus type and sensory modality) under conditions designed to promote the facilitated processing of specific unattended information compared to other information that was equally presented and also unattended. As mentioned in chapter three, Dewald and colleagues (2013) demonstrated the importance of temporal alignment for the facilitation of task-irrelevant information. These researchers utilized a compound RSVP stream of pictures with superimposed words and asked participants to respond to targets (i.e., immediate picture repetitions) while ignoring the superimposed words. Later, participants were given a surprise recognition test for the previously ignored words. Dewald et al. found that processing for explicitly presented, unattended, task-irrelevant words was facilitated when they were frequently paired with a task-relevant picture target that participants were attending and responding to, when compared to unattended words that were not aligned with a task-relevant picture, despite both word types being presented at an equal frequency.

6.1. Aims of this Dissertation

Thus, the current project adapted the experimental design from Dewald et al. (2013) with the primary objective of exploring the extent to which processing for unattended lexical and non-lexical stimuli may be facilitated under unimodal and cross-modal conditions, when participants are otherwise engaged in an attention-demanding task. In order to address this broad research question, this dissertation has two main aims:

- **Aim 1** was to evaluate the role of stimulus type (i.e., lexical vs. non-lexical information) in the facilitated processing of task-irrelevant stimuli
- **Aim 2** was to evaluate the role of multisensory presentation in the facilitated processing of task-irrelevant stimuli

These aims were addressed separately in two experiments. Experiment 1 focused on facilitated processing of task-irrelevant lexical and non-lexical items under unimodal conditions. Specifically, Experiment 1a presented participants with a compound visual stream containing lexical (i.e., written words) and non-lexical (i.e., pictures) information followed by a visual surprise recognition test for the unattended dimension of the stream (i.e., words or pictures, depending on condition). Experiment 1b utilized a compound auditory stream with lexical (i.e., auditory words) and non-lexical (i.e., common sounds) information, followed by an auditory surprise recognition test for the unattended items. Experiment 2 focused on facilitated processing of lexical and non-lexical items under cross-modal (i.e., multisensory) conditions. Specifically, Experiment 2a presented participants with a compound auditory/visual stream containing lexical and non-lexical information. Here, participants always ignored the *visual* dimension of the stream (i.e., written words or pictures, depending on condition), while attending to the auditory dimension, followed by a visual surprise recognition test. Experiment 2b, also utilized a compound auditory/visual stream, however in this experiment participants always ignored the *auditory* dimension of the compound stream (i.e., auditory words or common sounds, depending on condition), while attending to the visual dimension, followed by an auditory surprise recognition test. Taken together, this doctoral dissertation extends the findings of Dewald et al. (2013), by manipulating the sensory modality (i.e., auditory or visual) and the stimulus type (i.e., lexical or non-lexical) of the attended and ignored information.

7. Experiment 1: Unimodal Conditions

Experiment 1 focused on exploring the extent to which processing for explicitly ignored lexical and non-lexical information may be facilitated under unimodal conditions. Experiment 1a examined facilitated processing of lexical and non-lexical information in the visual modality, while Experiment 1b examined facilitated processing for lexical and non-lexical information in the auditory modality.

7.1. Experiment 1a: Unimodal Visual Conditions

All stimuli were presented in the visual modality. Participants were presented with a visual stream containing both pictures and words and were required to monitor one stimulus dimension (i.e., pictures or words in a between subjects design) for task-relevant targets (i.e., immediate repetitions) while ignoring the other. After completing this attention-demanding primary task, the extent to which the ignored items may have been processed were evaluated via a visual surprise recognition test (see stimuli and procedure for details).

7.1.1. Participants

All participants in this doctoral dissertation were recruited from undergraduate courses at the University of Hawai`i at Mānoa in exchange for course credit. Previous work using this paradigm (Dewald et al., 2011, 2013; Rees et al., 1999;) recruited sixteen, or fewer, participants. Because the current study intended to replicate and extend the findings from this seminal work, a much larger N was collected for greater reliability and statistical power. Accordingly, a total of 116, English speaking, young adults were recruited for Experiment 1a. All participants were naïve to the purpose of the experiment and were able to utilize devices designed to correct visual or auditory impairments. Of the 116 participants recruited, 63 (n = 63, *M* age = 20.1, 42 female) were randomly assigned to ignore non-lexical visual information (i.e., attend written words / ignore pictures condition) while the remaining 53 (n = 53, *M* age = 20.8, 34 female) were assigned to ignore lexical visual information (i.e., attend pictures / ignore written words condition).

Of the 63 assigned to the attend written words / ignore pictures condition, six participants were removed from data analysis due to having a high miss rate (i.e., 1*SD* above the mean), while an additional 10 were removed from data analysis for having a high false alarm (FA) rate (i.e., 1*SD* above the mean) during the primary task. While this criterion can be viewed as somewhat stringent, it was important that participants were adequately engaged with the primary task of detecting repetitions so as to ensure that attention did not stray to the ignored items throughout the duration of the experimental session. Of the 53 assigned to the attend pictures / ignore written words condition, eight participants were removed from data analysis due to having a high miss rate during the primary task, six for having a high FA rate during the primary task, and two for providing uniform responses (i.e., either all 'yes' or all 'no') during the surprise recognition test.

The reported analysis includes the remaining 47 participants in the attend written words / ignore pictures condition (n = 47, *M* age = 20.0, 30 female) and the remaining 37 participants in the attend pictures / ignore written words condition (n = 37, *M* age = 20.4, 24 female). All participants were presented with informed consent (See appendix B) prior to beginning the experiment and debriefed upon completion (See appendix C).

7.1.2. Stimuli

A total of 50 pictures (average size of five – ten centimeters) were selected from the Snodgrass and Vanderwart (1980) picture database (see Appendix D). Each picture was randomly rotated +/-30 degrees from its original orientation to ensure that the primary task was sufficiently demanding in each version of the experiment (see Rees et al., 1999; Sinnett et al., 2006). A total of 50 high frequency English words were selected from the MRC psycholinguistic database (Wilson, 1988, see Appendix D). Consistent with previous research using these same stimuli (Dewald et al., 2013; Walker et al., 2014; Walker et al., 2017), the selected words had an average length of five letters with a range of four – six letters and a frequency of 361 per million with a range of 135-782 per million. Each of the 50 pictures were superimposed with one of the 50 high frequency English words (precise stimulus combinations varied depending on experimental condition – see below for details). Care was taken to ensure that picture-word combinations did not have any semantic relationship. The words were superimposed over the rotated pictures in bold, capitalized, letters and presented in Arial font (24 points, subtending 0.84 centimeters). For the primary task, a stream of 960 combined picture-word items (height and width not exceeding ten centimeters) were created and presented in a RSVP stream using DMDX software (Forster & Forester, 2003).

7.1.2.1. Attended Written Words / Ignored Pictures

For this condition, participants attended lexical visual information (i.e., written words) during the primary task and were tested on the ignored non-lexical visual information (i.e., pictures) during the surprise recognition test. Therefore, all 50 words (Wilson, 1988) were presented during the primary task superimposed over one of eight randomly selected pictures from the previously obtained set of 50 (Snodgrass & Vanderwart, 1980). Immediately repeated words in the RSVP stream served as the task-relevant targets in the primary task. The RSVP stream was broken into eight blocks of 120 trials. The presentation was pseudorandomized such that in each block an immediate word repetition occurred, on average, in of one out of every eight trials. This created a mean of 15 task-relevant targets (word repetitions) per block, resulting in a total of 120 trials of exposure to a task-relevant target (word repetition) and a simultaneously presented task-irrelevant picture throughout the experiment.

Of the eight pictures that were selected and superimposed with words in the 960 trial RSVP stream, one was randomly selected to appear in temporal alignment (i.e., target-aligned or TA) with every task-relevant target (i.e., immediate word repetition). All other pictures appeared with the same total frequency, but were superimposed with words that did not immediately repeat (i.e., non-aligned or NA). In other words, a single picture was selected and always paired with the presentation of an immediate target word repetition, however all pictures were presented an equal number of times (i.e., 120 presentations) during the exposure stage. Eight iterations of this condition were created for which each of the eight pictures served as the image that was aligned with the target word repetitions. To control for any possible differences that may exist with regard to individual picture saliency, the presentation was randomized between participants. This manipulation was included to replicate the dependent measure and parallel the quantity of items and exposure to irrelevant stimuli employed by Dewald and colleagues (Dewald & Sinnett, 2012; Dewald et al., 2013; see also Watanabe et al., 2001).

7.1.2.2. Attended Pictures / Ignored Written Words

The exact same experimental construction was employed for the condition in which participants attended the non-lexical visual information (i.e., pictures) during the primary task and were tested on ignored lexical visual information (i.e., written words) during the surprise recognition test. Here, all 50 pictures (Snodgrass & Vanderwart, 1980) that were presented during the primary task were superimposed with one of eight randomly selected words from the previously obtained set of 50 (Wilson, 1988). Immediate picture repetitions served as the task-relevant target while a single word was selected and always paired with the presentation of an immediate target picture repetition (TA). As before, all other words appeared superimposed with non-repeating pictures (NA) in the RSVP stream, with all words presented an equal number of times (i.e., 120 presentations). Again, to control for possible differences in individual word saliency, eight versions were created in which each word served as the TA word and presentation was randomized between participants.

7.1.2.3. Surprise Recognition Test

The surprise recognition tests consisted of a total of either sixteen pictures or words (depending on the condition of the exposure stage). For those in the attend written words / ignore pictures condition, participants were tested on the previously ignored pictures, while those in the attend pictures / ignore written words condition were tested on the previously ignored words. For each version of the surprise recognition test, eight items (pictures or words) came from the previously viewed visual stream (the seven NA items and the single TA item), while the other eight consisted of never before seen foil items. The foil items were randomly selected from the remaining 42 items (pictures or words) of the originally selected set of 50 (Snodgrass & Vanderwart, 1980; Wilson, 1988) that were not used during the primary task in the experiment (recall that for each version of the experiment, only eight of the 50 items were used for the unattended stimulus dimension).

The recognition tests were randomized and presented using DMDX software (Forster & Forster, 2003). The pictures were presented in the same manner as during the primary task, but without words superimposed on top, and remained on the screen until a response was made. The words were presented in the same manner as during the primary task, but without pictures superimposed, in bold Arial font at 24 point, and remained on the screen until a response was made.

7.1.3. Procedure

Depending on the assigned condition (i.e., attend written words / ignore pictures or attend pictures / ignore written words) participants were asked to attend to one aspect of the visual stream (words or pictures), while ignoring the other. They were instructed to respond when they saw a target item (word or picture) immediately repeat in the RSVP stream by clicking the left mouse button with their preferred hand. Each item in the picture-word presentation was presented for 350ms with a 150ms inter-stimulus interval (ISI; blank screen) between each item for a stimulus onset asynchrony (SOA) of 500ms (see Figure 9). Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task (verified by experimenter observation and verbal confirmation of the participant). Immediately after the primary task, the surprise recognition test for the ignored items was administered. Participants were instructed to press the "B" key if they recalled having seen the picture (or word) during the primary task or, instead, the "V" key if they did not recall having seen the picture (or word) before. Response keys were counterbalanced across participants. The entire experimental session lasted for approximately 20-30 minutes.

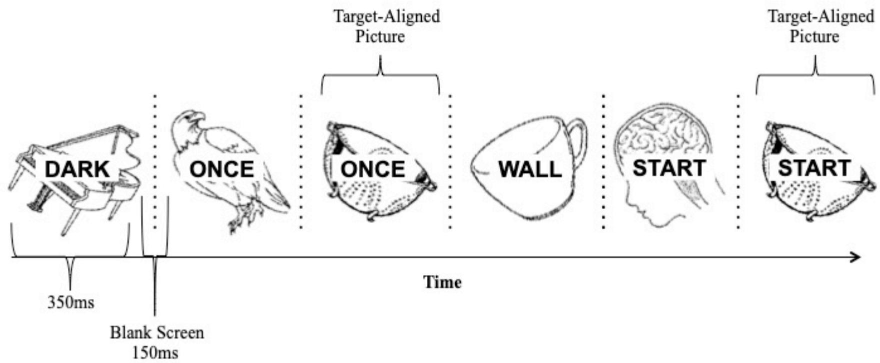


Figure 9. Schematic representation of the primary task, presented in the visual modality, in the attend words / ignore pictures condition. Immediately repeated words (i.e., 1-back task, Kirchner, 1958) served as the target in the identification task (e.g., "once" and "start") while superimposed images were the ignored items. Pictures that appeared with immediate word repetitions were the TA items (e.g., the colander); all other pictures were NA items. Notice that the TA item was always the same. However, all ignored items were presented an equal number of times during the entirety of the primary task.

7.1.4. Predictions

Regarding primary task performance, significant differences in performance rates (as indicated by d' scores and reaction time) were not anticipated between attended written words and attended pictures. All items presented during the primary task are highly salient, easy to identify, and familiar for this subject group. Therefore, participants should have little difficulty identifying targets in the RSVP stream. Based on previous findings by Carr et al. (1982), which demonstrated that participants were faster to categorize pictures compared to words, it was predicted that participants would have faster RTs when identifying picture repetitions compared to word repetitions during the primary task. Note, while this is interesting, the critical analyses for the questions being asked in this doctoral dissertation pertain to the surprise recognition task.

Regarding the surprise recognition test, previous investigations conducted by Watanabe and colleagues (Seitz & Watanabe, 2003; Watanabe et al., 2001) demonstrated a facilitatory effect of temporal alignment for task-irrelevant stimuli. The findings presented by Watanabe et al. (2001, see also Seitz & Watanabe, 2003) using relatively simple stimuli (i.e., motion detection) were later replicated by Dewald et al. (2013) using the exact same complex stimuli (i.e., words and pictures) in the exact same paradigm as presented in this dissertation (see chapter three). Therefore, it was expected that, for both conditions (ignored pictures and ignored words) recognition rates for TA items would be significantly higher compared to NA items, regardless of stimulus type. This would result in a main effect for target-alignment, suggesting that temporal and spatial alignment between an ignored item and an attended target (i.e., an item requiring successful identification and subsequent execution of response operations) leads to facilitated processing of the ignored information. This facilitated processing for TA items was expected to produce higher recognition rates when that information was encountered again during the surprise recognition test, compared to equally presented unattended items that were not paired with an attended task target (i.e., NA items), regardless of stimulus type.

Predicting the extent to which stimulus type has a significant impact on the facilitation of task-irrelevant information in the visual modality was more tenuous as this has not explicitly been explored in the scientific literature to date. Given the available information regarding the extent to which pictures and words are processed differently under conditions of directed attention (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980) and when these items are actively ignored (Tipper, 1985; Tipper & Driver, 1988), it was predicted that unattended pictures were more likely to be facilitated, compared to words, in the present paradigm. Recall that Carr et al. (1982) observed that picture targets were more likely to be facilitated by a preceding prime compared to word targets; participants were faster to *categorize* pictures compared to words; and they were faster to *name* words compared to pictures. Faster categorization for pictures suggest this stimulus type may have more direct access to relevant semantic information compared to words. Furthermore, findings from Tipper (1985, see also Tipper & Driver, 1988) suggested that semantic information may still be accessed for both pictures and words when these items were actively ignored. Therefore, it was inferred that if pictures maintain more direct access to semantic information under conditions of directed attention (thereby facilitating the extent to which those items may be processed) then this may also hold true for conditions in which these items are actively ignored. Thus, Experiment 1a was anticipated to yield a main effect for stimulus type, meaning that – regardless of temporal alignment (TA or NA) previously ignored pictures would be recognized more accurately during the surprise recognition test compared to previously ignored words.

The critical question that remains is the degree to which stimulus type and target alignment were likely to interact under these circumstances. Specifically, if pictures are indeed processed more extensively than words when these items are actively ignored, and if temporal alignment of an ignored item with an attended target facilitates processing of the unattended information, will pictures receive greater facilitation compared to words when these unattended items are temporally aligned with task-relevant targets? Based on the previous literature noted above, this outcome was deemed to be plausible given how these types of information are purportedly handled by our cognitive system.

7.1.5. Results

7.1.5.1. Primary Task

Performance on the primary task was evaluated by calculating the proportion of hits (defined as a correct response to a target within 1000ms of stimulus onset, see Ngo et al., 2011), FAs, and RTs for identified targets (i.e., immediate picture or written word repetitions in the RSVP stream) separately for both groups. Successful completion of the primary task was defined as obtaining a hit rate significantly above chance. During the primary task, a target appeared, on average, in one of every eight trials. Therefore, chance was calculated as the probability of obtaining a hit in any given presentation of eight trials (i.e., 12%). Analyses of primary task hit rate revealed that participants in both groups obtained a proportion of hits significantly above chance (12%) [attend written words / ignored pictures (i.e., attend written words): $M = 0.64$, $SE = 0.02$, $t(46) = 33.65$, $p <$

0.001; attend pictures / ignored written words (i.e., attend pictures): $M = 0.64$, $SE = 0.01$, $t(36) = 38.47$, $p < 0.001$].

In order to account for potential differences in target detection sensitivity that may occur when identifying different types of targets (i.e., written words or pictures), signal detection analysis (d') was performed using participant's hits and FAs during the primary task and comparing d' between groups. Participants in the attend written words condition had a significantly higher d' compared to those in the attend pictures condition [attend written words: $M = 3.55$, $SE = 0.09$ vs. attend pictures: $M = 2.97$, $SE = 0.05$, $t(82) = 5.10$, $p < 0.001$] (see Figure 10), suggesting a higher level of target detection sensitivity among those in the attend words condition compared to those in the attend pictures condition. All analyses met Bonferroni corrections for multiple comparisons ($p < 0.01$)[3].

[3] Bonferroni corrections for primary task performance included four main analyses (0.05/4 = 0.01): two t-tests on hit rates against chance, one between group t-test on d' rates, and one between group t-test on RTs.

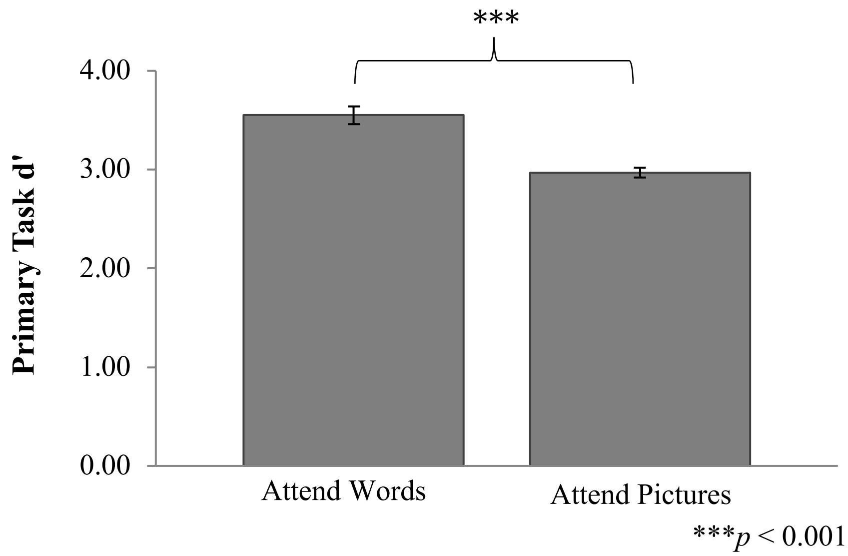


Figure 10. d' rates for target identification during the primary task in Experiment 1a. All stimuli were presented in the visual modality. "Attend Words" indicates the condition in which participants monitored the RSVP stream for immediate word repetitions and "Attend Pictures" indicates the condition in which participants monitored the RSVP stream for immediate picture repetitions. Error bars represent the standard error for each variable. Those in the attend words condition had a d' score significantly higher (M = 2.97, SE = 0.05) than those in the attend pictures condition (M = 3.55, SE = 0.09, $p < 0.001$), suggesting higher sensitivity to target identification among those in the attend words condition.

Finally, in order to evaluate processing speed for each stimulus type (i.e., pictures or written words) presented during the primary task, participants' RTs to identified targets were aggregated and the mean RTs were compared between groups. There was no significant difference in RT to identified targets between groups [attend written words: M = 492ms, SE = 6.94 vs. attend pictures: M = 484ms, SE = 7.83, $t(82) = 0.72$, $p = 0.472$]

7.1.5.2. Surprise Recognition Test

Performance on the surprise recognition was evaluated by calculating the mean proportion of hits, FAs, and correct rejections (CRs), separately for both groups. During the surprise recognition test, a hit was defined as the successful identification of a previously seen item (i.e., correctly responding "yes" to TA or NA items); a miss was defined as incorrectly rejecting a previously seen item (i.e., incorrectly responding "no" to TA or NA items); a FA was defined as incorrectly confirming the presence of a foil item (i.e., incorrectly responding "yes" to a foil item); and a CR was defined as correctly rejecting a foil item (i.e., correctly responding "no" to a foil item).

Participants' performance on the surprise recognition test was first assessed by calculating an overall accuracy score (i.e., hits + CRs) for each group and comparing performance on this metric against chance. Successful completion of the surprise recognition test was defined as obtaining an accuracy score significantly above chance. The recognition tests contained half old (i.e., either TA or NA) and half new (i.e., foil) items, therefore, chance was defined as the probability of obtaining a hit on any given trial (i.e., 0.5). Both groups obtained an accuracy score significantly above chance (50%) [attend pictures / ignored written words (i.e., ignored written words): $M = 0.77$, $SE = 0.02$, $t(36) = 14.39$, $p < 0.001$; attend written words / ignored pictures (i.e., ignored pictures): $M = 0.89$, $SE = 0.02$, $t(46) = 23.86$, $p < 0.001$] (see Figure 11). The accuracy scores for both groups failed to meet the assumption of normality; therefore, these analyses were corroborated by One-Sample Wilcoxon Signed-Rank Tests [ignored written words: $Z = 5.30$, $p < 0.001$, ignored pictures: $Z = 5.99$, $p < 0.001$]. All analyses met Bonferroni corrections for multiple comparisons ($p < 0.02$)[4].

[4] Bonferroni corrections for overall accuracy performance during the surprise recognition test included three main analyses (0.05/3 = 0.02): two t-tests on accuracy scores (hits + CRs) against chance and one between-group t-test on d' rates.

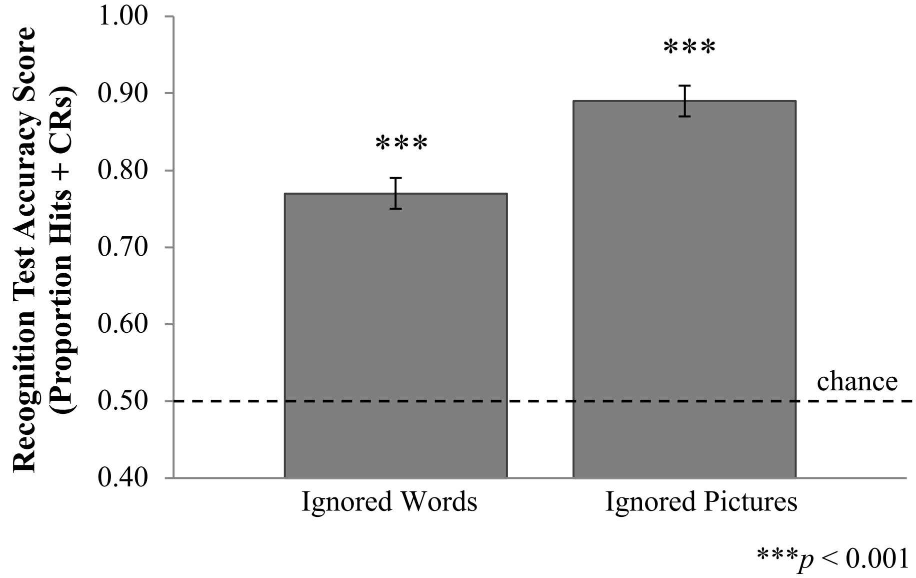


Figure 11. Accuracy scores (i.e., proportion of hits + CRs) for the surprise recognition test in Experiment 1a. All stimuli were presented in the visual modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored written words and "Ignored Pictures" indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. Both groups had an accuracy score significantly above chance (i.e., 0.50, $p < 0.001$), suggesting successful completion of the surprise recognition test.

In order to account for potential differences in target detection sensitivity that may occur when identifying different previously ignored items (i.e., written words or pictures), signal detection analysis (d') was performed. Participants in the ignored written words condition had a significantly lower d' compared to those in the ignored pictures condition [ignored written words: $M = 3.03$, $SE = 0.31$ vs. ignored pictures: $M = 5.18$, $SE = 0.37$, $t(82) = 4.31$, $p < 0.001$] (see Figure 12), suggesting a lower level of target detection sensitivity among those in the ignored written words condition compared to those in the ignored pictures condition. The d' variables for both groups failed to meet the assumption of normality; therefore, this analysis was corroborated by Mann-Whitney U Test [U = 4.21, $p < 0.001$]. Both analyses met the Bonferroni corrected significance level ($p < 0.02$).

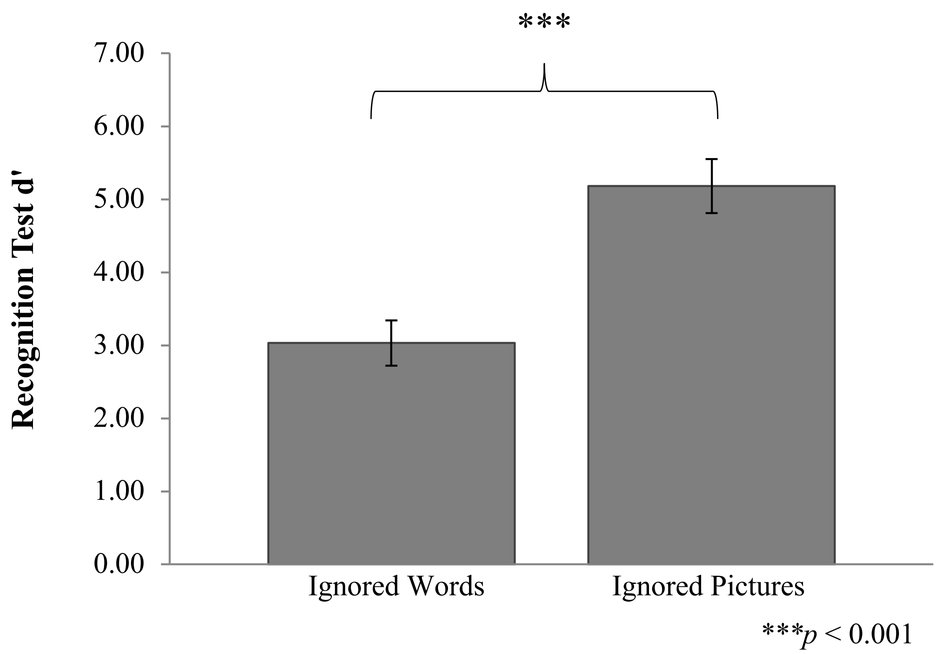


Figure 12. d' rates for target identification during the surprise recognition test in Experiment 1a. All stimuli were presented in the visual modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored written words and "Ignored Pictures" indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. Those in the ignored words condition had a d' score significantly lower ($M = 3.03$, $SE = 0.31$) than those in the ignored pictures condition ($M = 5.18$, $SE = 0.37$, $p < 0.001$), suggesting lower sensitivity to target identification among those in the ignored words condition during the surprise recognition test.

Regarding the critical analysis, a two-factor (2x2) ANOVA was conducted on recognition performance for the surprise recognition test, with focus of attention (ignored written words or ignored pictures) as the between subjects factor and target-alignment (TA or NA) as the within subjects factor. There was a main effect for target-alignment indicating that, overall, TA items [$M = 0.83$, $SE = 0.04$] were recognized significantly more often than NA items [$M = 0.72$, $SE = 0.02$, $F(1,82) = 7.12$, $p = 0.009$]. There was a main effect for group type indicating that ignored written words [$M = 0.64$, $SE = 0.03$] were recognized significantly less often than ignored pictures [$M = 0.82$, $SE = 0.02$, $F(1,82) = 11.02$, $p = 0.001$]. Finally, there was no interaction [$F(1,82) = 0.416$, $p = 0.521$], suggesting that the pattern of facilitation was comparable between groups (see Figure 13). Despite a lack of interaction, planned comparisons described below explored recognition rates for TA and NA items specifically within, and between, each condition (ignored written words or ignored pictures).

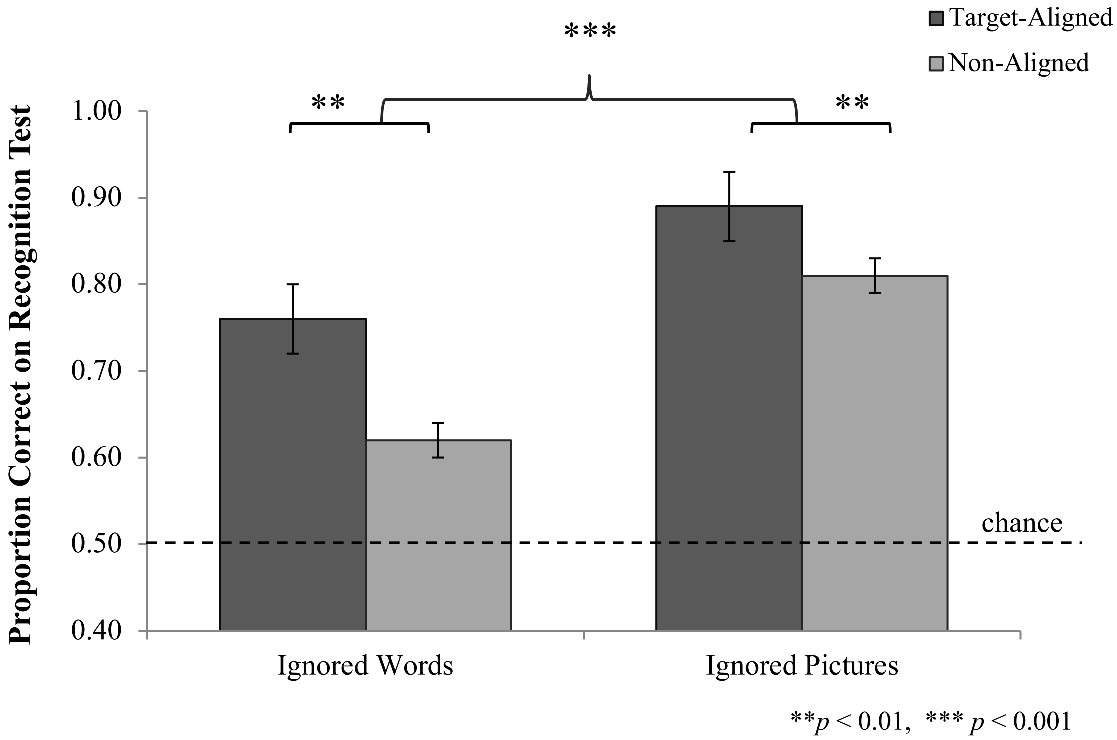


Figure 13. Main effects from the critical analysis of the surprise recognition test in Experiment 1a. All stimuli were presented in the visual modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored written words and "Ignored Pictures" indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. All items were recognized at rates significantly above chance (0.50 – significance not shown). There was a main effect for target-alignment (**$p < 0.01$), suggesting that TA items were recognized more often than NA items. There was also a main effect for stimulus type (***$p < 0.001$), suggesting that previously ignored pictures were recognized more often that previously ignored written words. There was no interaction.

7.1.5.2.1. Ignored Written Words

Overall, participants were able to recognize the previously ignored written words (TA and NA) statistically better than chance (50%) [$M = 0.64$, $SE = 0.03$, $t(36) = 4.24$, $p < 0.001$], which met Bonferroni corrections for multiple comparisons ($p < 0.005$)[5]. Recognition for TA words [$M = 0.76$, $SE = 0.07$, $t(36) = 3.59$, $p < 0.001$], and NA words [$M = 0.62$, $SE = 0.03$, $t(36) = 3.59$, $p < 0.001$] was each better than chance and met Bonferroni corrected significance levels ($p < 0.005$). Both variables failed to meet assumptions of normality; therefore, the TA word analysis was corroborated by a Binomial Test[6] [the observed proportion of TA words (0.76) was significantly higher than the expected proportion (0.50), $p = 0.003$] and the NA word analysis was corroborated by a One-Sample Wilcoxon Signed-Ranks Test [NA words: Z = 3.05, $p = 0.002$]. Finally, when compared to each other, TA words were recognized significantly more often than NA words according to conventional standards ($p < 0.05$) [$t(36) = 1.94$, $p = 0.03$], corroborated by Wilcoxon Signed Ranks Test [Z = 1.75, $p = 0.04$], however, both analyses failed to meet Bonferroni corrected significance level ($p < 0.005$).

7.1.5.2.2. Ignored Pictures

Overall, participants were able to recognize the previously ignored pictures (TA and NA) statistically better than chance (50%) [$M = 0.82$, $SE = 0.02$, $t(46) = 12.98$, $p < 0.001$], which met Bonferroni corrections for multiple comparisons ($p < 0.005$). Recognition for TA pictures [$M = 0.89$, $SE = 0.05$, $t(46) = 8.66$, $p < 0.001$], and NA pictures [$M = 0.81$, $SE = 0.03$, $t(46) = 11.79$, $p < 0.001$] was each better than chance and met the Bonferroni corrected significance level ($p < 0.005$). Both variables failed to meet assumptions of normality; therefore, the TA picture analysis was corroborated by a Binomial Test [the observed proportion of TA pictures (0.89) was significantly higher than the expected proportion (0.50), $p < 0.001$] and the NA pictures analysis was corroborated by a One-Sample Wilcoxon Signed-Rank Test [NA pictures: Z = 5.78, $p < 0.001$]. Finally, when compared to each other, TA pictures were recognized significantly more often than NA pictures according to conventional standards ($p < 0.05$) [$t(46) = 1.77$, $p = 0.04$], corroborated by Wilcoxon Signed Ranks Test [Z = 2.36, p = 0.009], however both analyses failed to meet Bonferroni corrected significance level ($p < 0.005$).

---

[5] Bonferroni corrections for post hoc investigations included ten main analyses (0.05/10 = 0.005): four t-tests on TA and NA items for the ignored words condition, four t-test on TA and NA items for the ignored pictures condition, and two t-test on TA and NA items between conditions.

[6] A Binominal test was used on TA items because this variable is binary.

7.1.5.2.3. Analysis by Target-Alignment

Comparing performance on the surprise recognition test for TA and NA items between each group revealed that there was no significant difference in recognition rates for TA items between the ignored written words condition [$M = 0.76$, $SE = 0.07$] and ignored pictures condition [$M = 0.89$, $SE = 0.05$, $t(82) = 1.69$, $p = 0.097$], corroborated by Mann-Whitney U Test [U = 1.66, $p = 0.097$]. However, NA items for ignored written words [$M = 0.62$, $SE = 0.03$] were recognized significantly less often than NA items for ignored pictures [$M = 0.81$, $SE = 0.03$, $t(82) = 4.61$, $p < 0.001$] (see Figure 14), corroborated by Mann-Whitney U Test [U = 4.17, $p < 0.001$] both of which met the Bonferroni corrected significance level ($p < 0.005$).

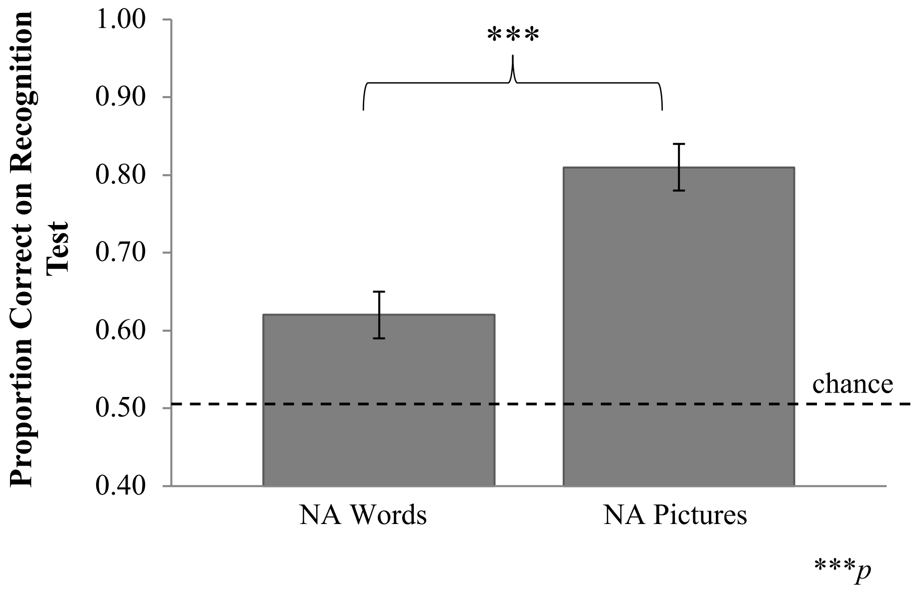


Figure 14. Recognition rates for NA items during the surprise recognition test in Experiment 1a. All stimuli were presented in the visual modality. "NA Words" indicates recognition rates for non-aligned written words from the "Ignored Words" condition and "NA Pictures" indicates recognition rates for non-aligned pictures from the "Ignored Pictures" condition. Error bars represent the standard error for each variable. Both NA items were recognized at rates significantly above chance (0.50 – significance not shown). Those in the ignored words condition recognized significantly fewer NA items (M = 0.62, SE = 0.03) than those in the ignored pictures condition (M = 0.81, SE = 0.03, $p < 0.001$), suggesting that NA words were facilitated to a lesser extent compared to NA pictures.

7.1.6. Discussion: Experiment 1a

Experiment 1a evaluated potential differences in the facilitated processing of previously unattended lexical (i.e., words) and non-lexical (i.e., pictures) stimuli under unimodal visual conditions in an attention-demanding task. In order to assess whether processing rates were modulated by stimulus type in the visual modality, between group comparisons were made on primary task and surprise recognition test performance. Regarding primary task performance, as predicted, both groups maintained a proportion of hits significantly above chance (i.e., 12%), which suggests that participants in both conditions were able to successfully identify targets in the compound visual stream. Interestingly, those in the attend written words condition had significantly higher d' scores compared to those in the attend pictures condition ($p < 0.001$), suggesting less sensitivity for signal detection among those monitoring the RSVP stream for picture repetitions compared to word repetitions. Thus, participants were more likely to respond to the attended pictures even when a target was not present. While somewhat surprising, this finding may suggest that response operations are more readily deployed when attending to pictures compared to words.

Regarding primary task reaction times, it was predicted that participants in the attend pictures condition would have significantly faster RTs compared to those in the attend words condition. This prediction was made based on previous research by Carr and colleagues (1982), who found that participants were faster to categorize pictures compared to words. The lack of any significant difference here may be attributed to the fact that the primary task in the current experiment is quite different from that of Carr et al. Specifically, in the current experiment, participants did not categorize items, but instead were required to recognize immediate repetitions in the RSVP of words and pictures. Previous research suggests that recognition precedes categorization, but does not necessitate it (Grill-Spector & Kanwisher, 2005; Humphreys & Forde, 2001). Therefore, it is possible that word recognition and picture recognition may occur at similar speeds, while categorization for words takes longer than it does for pictures. It should be noted that, while findings from the primary task are interesting, the focus of this dissertation is on the critical analyses of the surprise recognition test.

When analyzing the surprise recognition test for previously unattended items (i.e., words or pictures), as predicted, both groups displayed overall accuracy scores (i.e., hits + CRs) that were significantly above chance. This finding suggests that both groups were successful at distinguishing between previously seen items and foil items. Furthermore, as predicted, participants who were tested on previously ignored pictures (i.e., ignored pictures condition) had a significantly higher d' score compared to participants who were tested on previously ignored words (i.e., ignored words condition). This finding suggests that those in the ignored pictures condition were better able to distinguish between previously seen pictures and foil pictures compared to those in the ignored words condition. Such a finding aligns well with the idea that pictures may be processed more readily than words, even when ignored, leading to higher recognition rates – in general – during the surprise recognition test compared to words.

Considering the critical analysis on the surprise recognition test data, a two-factor (2x2) repeated measures ANOVA was conducted on participant responses with target-alignment (TA vs. NA) as the within subjects factor and stimulus type (ignored pictures vs. ignored words) as the between subjects factor. Watanabe and colleagues (2001; see also Seitz & Watanabe, 2003) have demonstrated the facilitatory effect of temporal alignment for task-irrelevant stimuli and these findings were later replicated by Dewald et al. (2013) using the exact same complex stimuli (i.e., words and pictures) in the exact same paradigm as presented in the current experiment. These findings were again replicated here, as there was a main effect for target-alignment, suggesting that, overall, TA items (words and pictures) were recognized significantly more often than NA items (words and pictures) during the surprise recognition test, despite the fact that all unattended items were presented in equal frequency during the primary task. Thus, it appears that temporal and spatial alignment between an ignored item and an attended target (i.e., an item requiring successful identification and subsequent execution of response operations) leads to facilitated processing of the ignored information in the visual modality resulting in higher recognition rates when that information is encountered again at a later time, compared to equally presented unattended items that were not paired with an attended task target (i.e., NA items), regardless of stimulus type (words or pictures).

Next, the ANOVA revealed a main effect for stimulus type, indicating that previously ignored pictures (TA and NA) were recognized significantly more often than previously ignored words (TA and NA). This finding aligns with Tipper (1985, see also Tipper & Driver, 1988), who suggest that unattended words and pictures may be processed somewhat extensively, resulting in the semantic evaluation of both stimulus types, even when actively ignored. Furthermore, earlier research also suggests that *attended* pictures maintain more direct access to relevant semantic information compared to words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980). Thus, it appears that while pictures and words may both be processed somewhat deeply when actively ignored (Tipper, 1985), pictures appear to be processed more extensively; leading to higher recognition rates during the surprise recognition test compared to previously ignored words.

Regarding the interaction, it was difficult to predict with certainty whether or not target-alignment and stimulus type would interact in the current experiment. Given the available literature, an interaction between these two dimensions did seem plausible given how these types of information are purportedly handled by our cognitive system, though to date this had not been explicitly tested. However, the ANOVA revealed no significant interaction between these two dimensions. Thus, while TA items were, indeed, recognized more often than NA items, and previously ignored pictures were, indeed, recognized more often that previously ignored words, the effect of target-alignment remained constant between the two stimulus types. This finding suggests that the observed facilitation for TA items proceeds in a similar manner between unattended lexical and non-lexical stimuli, despite pictures being processed more readily compared to words.

Finally, despite the fact that an interaction was not observed, because the nature of this dissertation contains exploratory comparisons, pre-planned analyses were conducted to investigate the relationship between recognition rates for TA and NA items within and between each stimulus condition. For both stimulus types (words and pictures), recognition rates for TA and NA items were significantly above chance. Compared to each other, for both stimulus types, TA items were recognized more often than NA items under the conventional statistical criterion, and also confirmed by non-parametric analyses, although not when corrected for multiple comparisons. When comparing performance on recognition rates for TA and NA items *between* stimulus types, there was no significant difference in recognition for TA items between pictures and words; however, NA pictures were recognized significantly more often than NA words.

Seitz and Watanabe (2003; see also Dewald et al., 2013) hypothesized that processing for ignored TA items is facilitated to a greater extent compared to NA items by virtue of TA items appearing in temporal and spatial proximity with an item that was given directed attention and processed extensively (i.e., a target). While there was no significant difference in recognition rates between TA pictures and words, suggesting comparable rates of facilitation, NA pictures were recognized at rates significantly higher than NA words. Facilitated recognition for NA pictures, compared to NA words, suggests that unattended NA pictures may be processed more extensively compared to unattended NA words and this additional facilitation may be attributed to the nature of the stimulus type itself rather than the temporal and spatial presentation of the items in the RSVP stream as this dimension was held constant between unattended words and pictures.

Taken together, it seems that target-alignment plays a critical role in facilitated processing for unattended information (regardless of stimulus type) in the visual modality. Furthermore, non-lexical visual stimuli (pictures) appear to be facilitated more extensively compared to lexical visual stimuli (words) when actively ignored.

7.2. Experiment 1b: Unimodal Auditory Conditions

Experiment 1b focused on exploring the extent to which processing for explicitly ignored lexical and non-lexical information may be facilitated in the auditory modality. Participants were presented with a complex auditory stream containing both sounds (i.e., non-lexical information) and auditory words (i.e., lexical information). They were asked to monitor one stimulus dimension (i.e., sounds or auditory words) for task-relevant targets while ignoring the other. After completing this attention-demanding primary task, the extent to which the ignored items may have been processed were evaluated via a surprise recognition test (see stimuli and procedure for details).

7.2.1. Participants

A total of 102, naïve, English speaking, young adults were recruited in the same manner as that of Experiment 1a (see section 7.1.1 for details) and were able to utilize devices designed to correct visual or auditory impairments. Of the 102 participants recruited, 50 (n = 50, *M* age = 20.6, 30 female) were randomly assigned to ignore non-lexical auditory information (i.e., attend auditory words / ignore sounds condition) while the remaining 52 (n = 52, *M* age = 20.0, 35 female) were assigned to ignore lexical auditory information (i.e., attend sounds / ignore auditory words condition).

Of the 50 assigned to the attend auditory words / ignore sounds condition, eight participants were removed from the analysis due to having a high miss rate (i.e., 1*SD* above the mean), while an additional three were removed from data analysis for having a high FA rate (i.e., 1*SD* above the mean) during the primary task. Again, this stringent criterion was applied to ensure that participants were actively engaged in the primary task. Finally, one participant was removed for providing uniform responses (i.e., all 'yes') during the surprise recognition test. Of the 52 assigned to the attend sounds / ignore auditory words condition, nine participants were removed from data analysis due to having a high miss rate during the primary task, four were removed for having a high FA rate during the primary task, and three were removed for providing uniform responses (i.e., all 'yes' or all 'no') during the surprise recognition test.

The reported analysis includes the remaining 38 participants in the attend auditory words / ignore sounds condition (n = 38, *M* age = 20.8, 22 female) and the remaining 36 participants in the attend sounds / ignore auditory words condition (n = 36, *M* age = 20.3, 23 female). All participants were presented with informed consent (see appendix B) prior to beginning the experiment and debriefed upon completion (see appendix C).

7.2.2. Stimuli

The same 50 high-frequency English words (Wilson, 1988) from Experiment 1a were used, thus target-aligned and non-aligned words were identical to those utilized in Experiment 1a. To translate the conditions of the experiment to the auditory modality, a native English speaker's voice was recorded reading each selected word three times, after which three blind listeners chose the best exemplar of each auditory word (a fourth listener was recruited in order to break a three-way tie when needed). The selected recordings were edited to have the same length of presentation (350ms) and average amplitude. A total of 50 sound stimuli were extracted from a standard database of 100 familiar sounds and were edited to 350ms and similar average amplitude (see Sinnett et al., 2006). All auditory stimuli were presented binaurally to participants via headphones. All participants were able to adjust the volume to a comfortable level.

7.2.2.1. Attended Auditory Words / Ignored Sounds Condition

The same 50 auditory English words were presented with eight selected sounds from the 50 that were extracted from the standard sound database (see Sinnett et al., 2006, see Appendix D). Immediate auditory word repetitions in the compound auditory stream were always paired with one of the eight selected sounds (i.e., the TA sound), while the remaining seven sounds were presented with non-repeating words (i.e., NA sounds). All sounds were presented an equal number of times and eight versions of the experiment were created in which each of the eight sounds served as the TA item (see section 7.1.2.1 for details from the complementary visual condition).

7.2.2.2. Attended Sounds / Ignored Auditory Word Condition

The 50 sounds were presented with the eight selected auditory words from the 50 that were extracted from the MRS psycholinguistic database (Wilson, 1988). Immediate sound repetitions in the compound auditory stream were always paired with one of the eight selected auditory words (i.e., the TA word), while the remaining seven auditory words were presented with non-repeating sounds (i.e., NA words). All auditory words were presented an equal number of times and eight versions of the experiment were created in which each of the eight auditory words served as the TA item (see section 7.1.2.2. for details from the complementary visual condition).

7.2.2.3. Surprise Recognition Test

The surprise recognition test was created and presented in a manner nearly identical to that of Experiment 1a (see section 7.1.2.3. for details) with the critical difference being that stimuli were now presented in the auditory modality according to the same specifications outlined above.

7.2.3. Procedure

As in Experiment 1a, participants were asked to attend to either the sounds or auditory words presented in the auditory stream and respond to immediate target repetitions by pressing the left mouse button with their preferred hand. Each item in the sound-auditory word stream was presented for 350ms with a 150ms inter-stimulus interval (ISI; silence) for a stimulus onset asynchrony (SOA) of 500ms (i.e., identical to Experiment 1a, see Figure 15). Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task (verified by experimenter observation and verbal confirmation of the participant). As in Experiment 1a, the surprise recognition test was administered to all participants immediately after the primary task. Participants were instructed to press the "B" key if they had heard the item (sound or auditory word) during the primary task or, instead, the "V" key if they had not heard the item before (again, counterbalanced across participants).

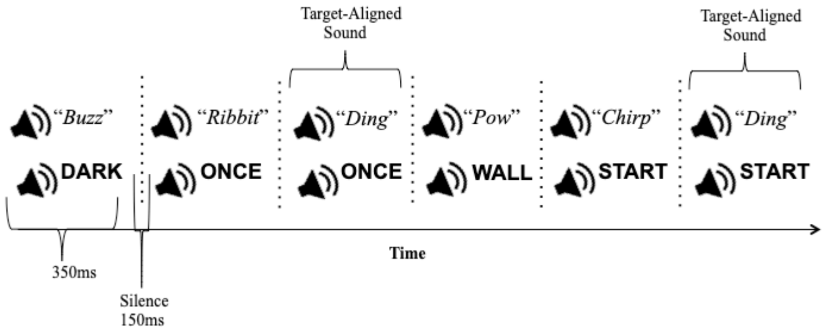


Figure 15. Schematic representation of the primary task, presented in the auditory modality, in the attended auditory words / ignored sounds condition. All stimuli were presented binaurally over headphones. In the schematic, auditory words are represented in bold, capitalized, letters, along the bottom row, while sounds are represented in italicized letters and presented in quotation marks along the top row. Immediate repetitions in the auditory word stream served as the target in the identification task (e.g., "once" and "start") while the sounds were the ignored items. Sounds that appeared with immediate auditory word repetitions were the TA items (e.g., "*Ding*"); all other sounds were NA items. Notice that the TA item was always the same. However, all ignored items were presented an equal number of times during the entirety of the primary task.

7.2.4. Predictions

Considering primary task performance, as with Experiment 1a, above chance performance was expected for both groups. Furthermore, significant differences in d' rates between groups (i.e., sounds vs. auditory words) were not expected. As with the visual modality, all auditory stimuli presented during the primary task are highly salient, easy to identify, and familiar for this subject group. Therefore, task performance was expected to be relatively high – meaning participants would have little difficulty identifying and responding to target repetitions. However, differences in RTs to identified targets were anticipated between groups. Recall from section 5.1.2 that Vitevitch and Luce (1998, 2016) found evidence to suggest that processing for sounds may be facilitated when semantic information is limited or absent. Arguably, the sounds selected for this experiment still maintain semantic information, which may override some of the facilitatory effects observed by Vitevitch and Luce (1998). However, as with pictures, sounds may still carry less semantic ambiguity compared to words, thus it was predicted that participants would respond faster to sound repetitions compared to auditory word repetitions during the primary task.

Next, when examining surprise recognition test performance, both groups were expected to demonstrate significantly above chance performance on their overall accuracy scores (i.e., hits + CRs). However, participants in the ignored sounds condition (i.e., those being tested on previously ignored sounds) were expected to have higher d' scores compared to those in the ignored auditory words condition (i.e., those being tested on previously ignored auditory words). This prediction was predicated on the assumption that non-lexical sounds have less semantic ambiguity and as such they may also be recognized more accurately – in general – during the surprise recognition test compared to auditory words. Thus, it was anticipated that participants would be better able to distinguish between previously heard items and newly presented foil items during the surprise recognition test for ignored sounds compared to ignored auditory words.

Based on findings from previous research on the facilitated processing of unattended visual items by target-alignment (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe, 2001) it was also anticipated that pairing an unattended auditory item (i.e., sound or auditory word) with an attended task-relevant auditory target would result in facilitated processing for the task-irrelevant information. This facilitated processing should then be reflected in the surprise recognition test, meaning that participants were expected to demonstrate overall higher recognition rates for TA items compared to NA items, regardless of stimulus type.

Given the paucity of previous research comparing the extent to which auditory words or sounds may be processed under conditions of directed attention or when they are explicitly ignored, it was, again, difficult to make precise predictions regarding which type of information may be more likely to undergo facilitated processing under the current conditions. Recall that Vitevitch and Luce (1998, 2016) found that processing for lexical information might be inhibited for words occurring in high-density neighborhoods compared to words occurring in low-density neighborhoods[7]. On the other hand, high-frequency, sub-lexical sounds did not appear to activate competitive representational nodes to the same extent, leading to facilitated processing when attention was directed toward this type of information (see Vitevitch & Luce, 1998). Based on this available information, a main effect for stimulus type was expected – specifically, participants should exhibit overall higher recognition rates for previously ignored sounds compared to previously ignored auditory words.

Finally, we must consider the extent to which the dimensions of stimulus type and temporal alignment might interact within the auditory modality. Previous studies investigating rates of inattentional deafness demonstrated that auditory lexical information was more likely to be processed when it shares featural characteristics with additional auditory lexical information that was being attended (Dalton & Frankel, 2012). Furthermore, many dichotic listening tasks have confirmed that ignored lexical information presented in the auditory modality may be evaluated – to some extent – for semantic content thereby allowing certain, high priority information (like one's name) to be facilitated even when attention was directed elsewhere (Cherry, 1953; Conway, Cowan, & Bunting, 2001; Dalton & Fraenkel, 2012; Giraudet, St-Louis, Scannella & Causse, 2015; Haykin & Chen, 2005; Macdonald & Lavie, 2011; Murphy & Greene, 2015; Raveh & Lavie, 2015; Treisman, 1960). However, if ignored auditory words are, indeed, evaluated at the semantic level, then activation of competitive representation nodes (i.e., akin to attended conditions observed by Vitevitch & Luce, 1998, 2016) may interfere with the facilitatory effects of temporal alignment for this stimulus type.

Investigations into processing for ignored, auditory, non-lexical information suggested that this type of information is subject to similar constraints as lexical auditory information (Koreimann et al., 2014). Recall, that Koreimann and colleagues found that, when attention was directed toward counting tympani beats in a musical piece, participants were highly likely to miss an additional electric guitar solo presented toward the end

[7] Luce and Pisoni (1998) define "high-density" words as those which sound similar to many other words (e.g., "rate") while "low-density" words are defined as those that sound similar to only a few other words (e.g., "bulb") in the representational system.

of the clip. However, sub-lexical auditory information appeared to be facilitated beyond its lexical counterpart when it was actively attended due to an apparent lack of competition from representational semantic equivalents (see Vitevitch & Luce, 1998, 2016). Therefore, unattended sounds may be more likely to be facilitated by temporal-alignment as this type of information appears to encounter less interference within our cognitive system (akin to non-lexical visual items – i.e., pictures). Based on these available studies, a significant interaction between stimulus type and temporal alignment was anticipated within the auditory modality. Specifically, while it was expected that all auditory TA items (lexical and non-lexical) would be facilitated compared to NA items, the magnitude of facilitated processing for unattended TA sounds was expected to be significantly above that of unattended TA auditory words. Following this same logic, unattended NA sounds would also be recognized at rates higher than those of previously unattended NA auditory words though whether or not this difference would reach a level of statistical significance was difficult to predict as the available literature is quite limited.

7.2.5. Results

7.2.5.1. Primary Task

Performance on the primary task was evaluated in the same manner as Experiment 1a (see section 7.1.5.1. for details). Analyses of primary task accuracy revealed that participants in both groups obtained a proportion of hits significantly above chance (12%) [attend auditory words / ignored sounds (i.e., attend auditory words): M = 0.72, SE = 0.01, $t(37) = 65.20$, $p < 0.001$; attend sounds / ignored auditory words (i.e., attend sounds): M = 0.63, SE = 0.01, $t(35) = 45.22$, $p < 0.001$].

Participants in the attend auditory words condition had significantly higher d' scores compared to those in the attend sounds condition [attend auditory words: M = 3.25, SE = 0.06 vs. attend sounds: M = 2.80, SE = 0.06, $t(72) = 5.51$, $p < 0.001$] (see Figure 16), suggesting a higher level of target detection sensitivity among those in the attend auditory words condition compared to those in the attend sounds condition. All analyses met Bonferroni corrections for multiple comparisons ($p < 0.01$)[8].

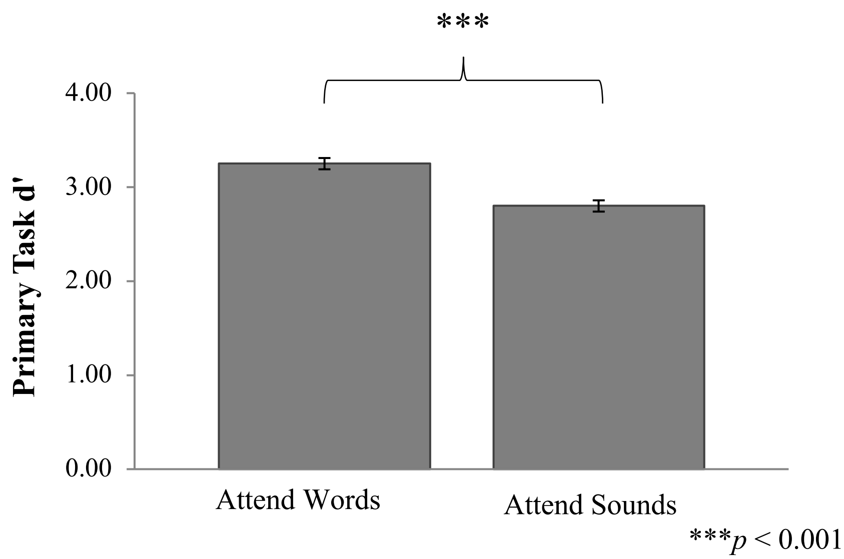


Figure 16. d' rates for target identification during the primary task in Experiment 1b. All stimuli were presented in the auditory modality. "Attend Words" indicates the condition in which participants monitored the RSAP stream for immediate word repetitions and "Attend Sounds" indicates the condition in which participants monitored the RSAP stream for immediate sound repetitions. Error bars represent the standard error for each variable. Those in the attend words condition had a d' score significantly higher (M = 3.25, SE = 0.06) than those in the attend sounds condition (M = 2.80, SE = 0.06, $p < 0.001$), suggesting higher sensitivity to target identification among those in the attend words condition.

Finally, in order to evaluate processing speed for each stimulus type (i.e. pictures or written words) presented during the primary task, participants' RTs to identified targets were aggregated and the mean RTs

[8] Bonferroni corrections for primary task performance included four main analyses (0.05/4 = 0.01): two t-tests on hit rates against chance, one between group t-test on d' rates, and one between group t-test on RTs.

were compared between groups. There was no significant difference in RT to identified targets between groups [attend auditory words: $M = 498$, $SE = 9.91$ vs. attend sounds: $M = 512$ms, $SE = 7.40$, $t(72) = 1.14$, $p = 0.257$].

7.2.5.2. Surprise Recognition Test

Performance on the surprise recognition was evaluated in the same manner as Experiment 1a (see section 7.1.5.2. for details). Participants in both groups obtained an accuracy score (Hits + CR) significantly above chance (50%) [attend sounds / ignored auditory words (i.e., ignored auditory words): $M = 0.73$, $SE = 0.02$, $t(35) = 11.49$, $p < 0.001$; attend auditory words / ignored sounds (i.e., ignored sounds): $M = 0.76$, $SE = 0.02$, $t(37) = 12.04$, $p < 0.001$] (see Figure 17), which met Bonferroni corrections for multiple comparisons ($p < 0.02$)[9].

[9] Bonferroni corrections for overall accuracy performance during the surprise recognition test included three main analyses (0.05/3 = 0.02): two t-tests on accuracy scores (hits + CRs) against chance and one between-group t-test on d' rates.

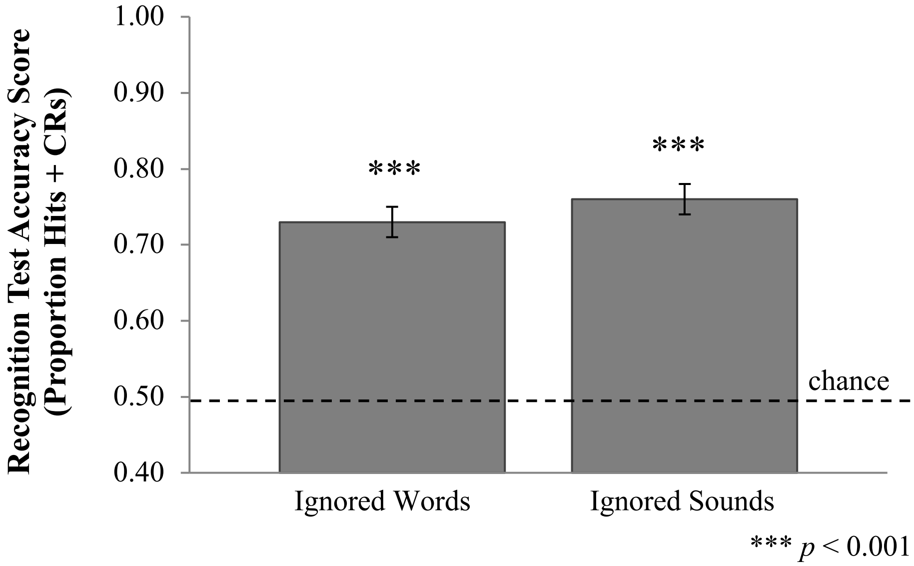


Figure 17. Accuracy scores (i.e., proportion of hits + CRs) for the surprise recognition test in Experiment 1b (auditory only). "Ignored Words" indicates the condition in which participants were tested on previously ignored auditory words and "Ignored Sounds" indicates the condition in which participants were tested on previously ignored sounds. Error bars represent the standard error for each variable. Both groups had an accuracy score significantly above chance (i.e., 0.50, $p < 0.001$), suggesting successful completion of the surprise recognition test.

There was no significant difference in d' rates between groups [ignored auditory words: $M = 2.16$, $SE = 0.34$ vs. ignored sounds: $M = 2.74$, $SE = 0.37$, $t(72) = 1.15$, $p = 0.256$], suggesting a comparable level of target detection sensitivity across groups. The d' data for both groups failed to meet the assumption of normality; therefore, this analysis was corroborated by a Mann-Whitney U Test [U = 0.97, $p = 0.335$].

Regarding the critical analysis, a two-factor (2x2) ANOVA was conducted on recognition performance for the surprise recognition test, with focus of attention (ignored auditory words or ignored sounds) as the between subjects factor and target-alignment (TA or NA) as the within subjects factor. There was no main effect for target-alignment indicating that, overall, TA items [$M = 0.58$, $SE = 0.06$] were not recognized significantly more often than NA items [$M = 0.68$, $SE = 0.02$, $F(1,72) = 2.71$, $p = 0.104$]. There was no main effect for group type indicating that ignored auditory words [$M = 0.64$, $SE = 0.03$] were not recognized at a rate significantly different from ignored sounds [$M = 0.70$, $SE = 0.03$, $F(1,72) = 0.04$, $p = 0.844$]. Finally, there was no interaction [$F(1,72) = 2.94$, $p = 0.09$] (see Figure 18). Despite a lack of interaction, planned comparisons described below explored recognition rates for TA and NA items specifically within, and between, each condition (ignored written words or ignored pictures).

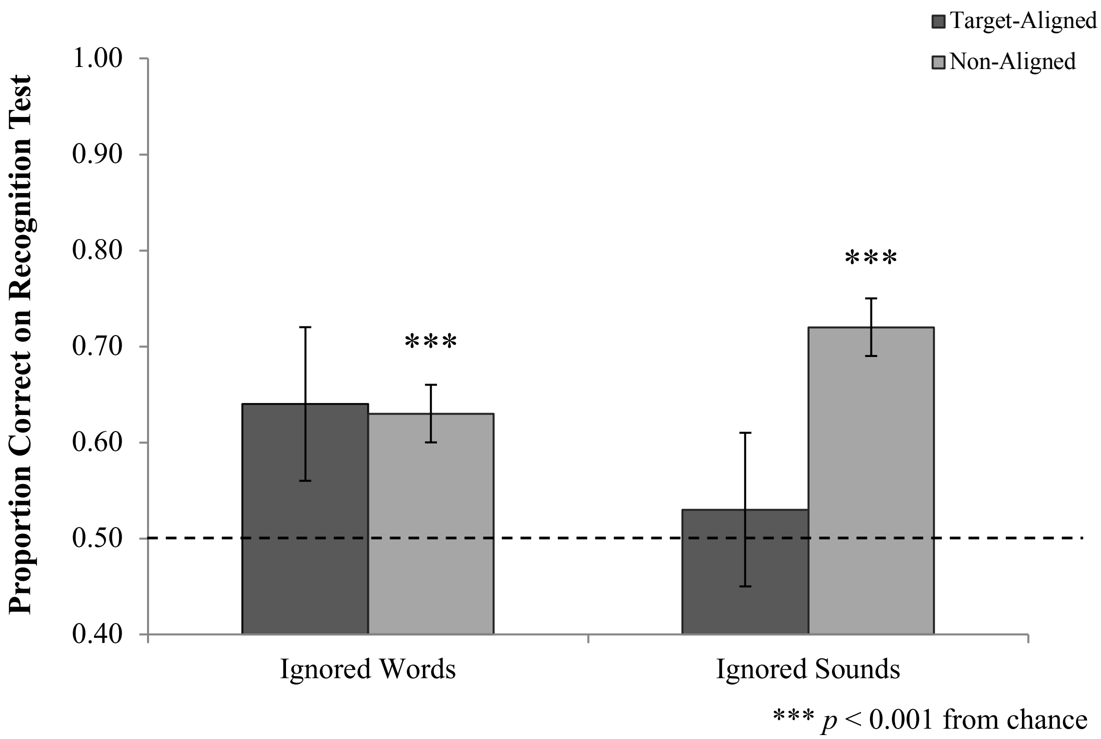


Figure 18. Results from the critical analysis of the surprise recognition test in Experiment 1b. All stimuli were presented in the auditory modality. “Ignored Words” indicates the condition in which participants were tested on previously ignored auditory words and “Ignored Sounds” indicates the condition in which participants were tested on previously ignored sounds. Error bars represent the standard error for each variable. There was no main effect for target-alignment, suggesting that TA items were not recognized more often than NA items. There was no main effect for stimulus type, suggesting that previously ignored sounds were not recognized more often that previously ignored auditory words. There was no interaction. NA items were recognized at rates significantly above chance (0.50, $p < 0.001$), while TA items were not.

7.2.5.2.1. Ignored Auditory Words

Overall, participants were able to recognize the previously ignored auditory words (TA and NA) statistically better than chance (50%) [$M = 0.64$, $SE = 0.03$, $t(35) = 4.78$, $p < 0.001$], which met Bonferroni corrections for multiple comparisons ($p < 0.005$)[10]. Recognition for TA words [$M = 0.64$, $SE = 0.08$, $t(35) = 1.71$, $p = 0.04$], and NA words [$M = 0.63$, $SE = 0.03$, $t(35) = 4.80$, $p < 0.001$] was each better than chance according to conventional significance levels ($p < 0.05$). The NA analysis met the Bonferroni corrected significance level ($p < 0.005$), while the TA analysis did not. Both variables failed to meet assumptions of normality; therefore, the TA word analysis was corroborated by a Binomial Test, which failed to reach significance [the observed proportion of TA words (0.64) was not significantly higher than the expected proportion (0.50), $p = 0.06$], while the NA word analysis was corroborated by a One-Sample Wilcoxon Signed-Ranks Test [NA words: Z = 3.90, $p < 0.001$]. Finally, when compared to each other, TA words were not recognized significantly more often than NA words [$t(35) = 0.05$, $p = 0.96$], corroborated by Wilcoxon Signed Ranks Test [Z = 0.40, $p = 0.697$].

7.2.5.2.2. Ignored Sounds

Overall, participants were able to recognize the previously ignored sounds (TA and NA) statistically better than chance (50%) [$M = 0.70$, $SE = 0.03$, $t(37) = 6.27$, $p < 0.001$], which met Bonferroni corrections for multiple comparisons ($p < 0.005$). Recognition for TA sounds was not significantly different from chance [$M = 0.53$, $SE = 0.08$, $t(37) = .32$, $p = 0.750$]. Recognition for NA sounds was better than chance [$M = 0.72$, $SE = 0.03$, $t(37) = 6.87$, $p < 0.001$] and met the Bonferroni corrected significant level ($p < 0.005$). Both variables failed to meet assumptions of normality; therefore, the TA sound analysis was corroborated by a Binomial Test [the observed proportion of TA sound (0.53) was not significantly different than the expected proportion (0.50), $p = 0.871$] and the NA sound analysis was corroborated by a One-Sample Wilcoxon Signed-Rank Test [NA sounds: Z = 4.62, $p < 0.001$]. Finally, when compared to each other, TA sounds were recognized significantly less often than NA sounds according to conventional standards [$t(37) = 2.34$, $p = 0.03$], corroborated by Wilcoxon Signed Ranks Test [Z = 2.38, $p = 0.02$], both of which failed to meet the Bonferroni corrected

[10] Bonferroni corrections for post hoc investigations included ten main analyses (0.05/10 = 0.005): four t-tests on TA and NA items for the ignored auditory words condition, four t-test on TA and NA items for the ignored sounds condition, and two t-test on TA and NA items between conditions.

significance level (p <0.005).

7.2.5.2.3. Analysis by Target-Alignment

Comparing performance on the surprise recognition test for TA and NA items between each group revealed that there was no significant difference in recognition rates for TA items between the ignored auditory words condition [M = 0.64, SE = 0.08] and ignored sounds condition [M = 0.53, SE = 0.08, $t(72)$ = 0.97, p = 0.33], corroborated by Mann-Whitney U Test [U = 0.97, p = 0.33]. However, NA items for ignored auditory words [M = 0.63, SE = 0.03] were recognized significantly less often than NA items for ignored sounds by conventional standards [M = 0.72, SE = 0.03, $t(72)$ = 2.02, p = 0.04], corroborated by Mann-Whitney U Test [U = 2.29, p = 0.02], however both analyses failed to meet the Bonferroni corrected significance level ($p < 0.005$).

7.2.6. Discussion: Experiment 1b

Experiment 1b evaluated potential differences in the facilitated processing of previously unattended lexical (i.e., words) and non-lexical (i.e., sounds) stimuli under unimodal auditory conditions in an attention-demanding task. In order to assess processing rates, between group comparisons were made on primary task and surprise recognition test performance.

Regarding primary task performance, as predicted, both groups maintained a proportion of hits significantly above chance (i.e., 12%), which suggests that participants in both conditions were able to successfully identify targets in the compound auditory stream. There was also no difference in d' rates between groups, suggesting comparable abilities to identify targets in the RASP between those in the attend auditory words condition and those in the attend sounds condition.

Next, it was predicted that those in the attend sounds condition would have significantly faster RTs to targets compared to those in the attend words condition. This was not the case, as there was no significant difference in RT between groups on the primary task. This prediction was made based on findings from Vitevitch and Luce (1998, 2016) who found evidence to suggest that processing speed for sounds are facilitated when semantic information is limited or absent. However, it is very likely that the sounds presented in the current experiment do contain semantic information, as these stimuli represent commonly occurring natural sounds such as horns, chimes, buzzes, and animal noises, and these semantic connections may be somewhat ambiguous (i.e., akin to semantic word processing). Thus, it is possible that the semantic nature of the sounds used here may override some of the facilitatory effects observed by Vitevitch and Luce (1998).

Regarding the surprise recognition test for previously unattended items (i.e., auditory words or sounds), as predicted, both groups displayed overall accuracy scores (i.e., hits + CRs) that were significantly above chance, and there was no significant difference in d' between groups. Taken together, these findings suggest that participants in both groups were able to successfully complete the surprise recognition test and they were comparable in their ability to distinguish between previously heard items (auditory words or sounds) and foil items, overall.

Considering the critical analysis on the surprise recognition test data, a two-factor (2x2) repeated measures ANOVA was conducted on participant responses with target-alignment (TA vs. NA) as the within subjects factor and stimulus type (ignored sounds vs. ignored auditory words) as the between subjects factor. Despite robust evidence for the role of temporal alignment in the facilitated processing of ignored information in the visual modality (Dewald et al., 2013; Seitz & Watanabe, 2003; Walker et al., 2014; Walker et al., 2017; Watanabe et al., 2001), there was no main effect for target-alignment here, suggesting that, overall, TA items (auditory words and sounds) were not recognized significantly more often than NA items (auditory words and sounds) during the surprise recognition test. Surprisingly, there was also no main effect for stimulus type, indicating that previously ignored sounds (TA and NA) were not recognized significantly more often than previously ignored auditory words (TA and NA). The interaction also failed to reach significance. While a lack of available literature makes interpreting these results somewhat cumbersome, there are some considerations worth exploring.

First, the primary task was presented under unimodal auditory conditions. This is important as the auditory modality lacks a level of stimulus specificity that the visual modality maintains. That is, simultaneously presented visual information may be segregated from one another more readily compared to simultaneously presented auditory information. Visual features such as shape, size, color (though not in the current experiment), and line serve as markers by which we can extract separate items from a compound visual stimulus with relative ease (Bundesen & Pedersen, 1983; Wagemans, Elder, Kubovy et al., 2012; Wolfe, 1994; Treisman, 1982). Additionally, corresponding segregation of the visual neural pathways have been noted in primate models (Livingstone, & Hubel, 1988; Self, van Kerkoerle, Super, & Roelfsema, 2013; Shipp & Zeki, 1985) and visual segregation has been demonstrated to interact with attention (Gilbert & Li, 2013; Morey, 2018; Poort, Raudies, Wannig, Lamme, Neumann, & Roelfsema, 2012) thereby guiding the visual perceptual representation. However, auditory information lacks some of the distinct featural specificity found in the visual modality, meaning it relies much more heavily on temporal markers, making it more difficult to segregate concurrently presented sounds from one another. Specific properties of auditory perceptual segregation include sensitivity to dimensions such as: acoustic rhythm (Notter, Hanke, Murray, & Geiser, 2018; Snyder, 2015), variations in stimulus intensity (Henny Yeung, Bhatara, & Nazzi, 2018), temporal onset and offset of the stimuli (Simon & Winkler, 2018); temporal neural oscillation (Teki, Barascud, Picard, Payne, Griffiths, & Chait, 2016; Teng, X., Tian, Rowland, & Poeppel, 2017; Tóth, Kocsis, Háden, Szerafin, Shinn-Cunningham, & Winkler, 2016), and interactions with top-down control (Wang, Zhang, Zou, Luo, & Ding, 2018).

Importantly, Simon and Winkler (2018) noted that lowering the ISI between concurrently presented auditory stimuli increases the likelihood of the listener experiencing an integrated auditory percept. This is especially likely under the primary task conditions presented in Experiment 1b, as each individual signal (i.e., sound and word) in the compound auditory stream was matched for relative onset, duration, and amplitude, which serve as important markers for segregating the individual sound streams in a compound auditory stimulus. Thus, the very fact that the information was presented in the auditory modality, under carefully controlled conditions, may mitigate the facilitatory effects of temporal alignment observed in the visual modality. That is, participants may have had a difficult time segregating the auditory sound from the auditory word during the primary task, leading to some level of signal integration between the unattended and attended sound streams regardless of target-alignment. Arguably, this would also make the primary task quite difficult, which may have increased the amount of cognitive load (Cartwright-Finch & Lavie, 2007; Lavie, 2005, 2010) associated with the task, adding an additional processing constraint for the unattended portion of the auditory stream, further mitigating the effects of target-alignment.

Second, it was expected that previously ignored sounds would be recognized significantly more often than previously ignored auditory words. This prediction was made based on findings from dichotic listening tasks, demonstrating that ignored lexical information presented in the auditory modality may be evaluated – at least to some extent – for semantic content (Cherry, 1953; Conway, Cowan, & Bunting, 2001; Dalton & Fraenkel, 2012; Giraudet, St-Louis, Scannella & Causse, 2015; Haykin & Chen, 2005; Macdonald & Lavie, 2011; Murphy & Greene, 2015; Raveh & Lavie, 2015; Treisman, 1960). Thus, if unattended words are evaluated for semantic content, and if accessing semantic associations for words leads to activation of competitive representation nodes within the semantic network (Vitevitch & Luce, 1998, 2016), then it seemed likely that these items would be processed to a lesser extent compared to unattended sounds. However, as stated earlier, the sounds selected for this experiment are very likely to contain semantic information, and the precise semantic associations carried by each sound may be somewhat ambiguous (compared to non-lexical visual items like pictures). Furthermore, little is known about the extent to which natural sounds may activate competitive representations within the semantic network, as words are believed to do. Though, it should be noted that at least one current study suggested that processing for ignored, auditory, non-lexical information may be subject to similar constraints as ignored, auditory, lexical information (Koreimann et al., 2014). Taken together, the findings suggest that both the unattended auditory words and the sounds used here were processed in a comparable manner during the primary task, resulting in similar overall recognition rates between stimulus types during the recognition test.

Finally, despite the fact that an interaction was not observed, pre-planned analyses explored the relationship between recognition rates for TA and NA items within and between each stimulus condition. For those in the ignored auditory word condition, recognition rates for TA words and NA words were both significantly above chance according to conventional standards, though only recognition rates for NA words met the Bonferroni criterion for multiple comparisons. There was also no significant difference in recognition rates between TA words and NA words. For those in the ignored sounds condition, recognition rates for TA sounds were not significantly different from chance, while NA sounds were, also meeting Bonferroni corrected significance levels. Furthermore, TA sounds were recognized significantly less often than NA sounds according to conventional levels of significance, but not when corrected for multiple comparisons. Finally, comparing performance on recognition rates for TA and NA items *between* stimulus types revealed no significant difference in recognition for TA items between previously ignored sounds and auditory words, however, NA sounds were recognized significantly more often than NA auditory words by conventional standards, though this failed to meet Bonferroni criterion for multiple comparisons. Thus, it appears that unattended TA auditory items were not likely to be facilitated to levels above chance recognition, while NA auditory items were, however it also seems unlikely that target-alignment played a critical role under the current conditions.

This finding was surprising, as previous work on target-alignment under these conditions in the visual modality predicted the opposite effect (Dewald et al., 2013, Seitz & Watanabe, 2003; Watanabe et al., 2001). As mentioned before, the difficult nature of extracting individual signals from a compound auditory stimulus likely increased cognitive load for the primary task (Cartwright-Finch & Lavie, 2007; Lavie, 2005, 2010). The added dimension of identifying an attended task target may have further increased processing constraints such that task-irrelevant auditory information appearing with a task-relevant auditory target underwent some level of inhibited processing, compared to NA items, in order to free up the necessary resources required for task completion. While the extent to which cognitive load operates the same in the auditory modality as it does in the visual modality is yet unclear, mixed evidence has suggested that similarities do exist (Murphy, Spence, & Dalton, 2017) and evidence from dichotic listening tasks demonstrated inhibited processing for unattended auditory streams (Cherry, 1953; Conway, Cowan, & Bunting, 2001; Dalton, & Fraenkel, 2012; Giraudet, St-Louis, Scannella, & Causse, 2015; Haykin & Chen, 2005; Macdonald & Lavie, 2011; Murphy & Greene, 2015; Raveh & Lavie, 2015; Treisman, 1960). Furthermore, inhibited processing for task-irrelevant information has also been reported using this type of paradigm with slightly different presentation parameters in the visual modality (Dewald & Sinnett, 2011a; 2011b; Dewald et al., 2011; Rees et al., 1999; Sinnett et al., 2006; Tsushima et al., 2008).

Taken together, it seems that target-alignment did not play a critical role in facilitated processing for unattended information (regardless of stimulus type) in the auditory modality. This may be attributed to the fact that presenting the primary task in the auditory modality resulted in higher rates of signal integration between the attended and ignored streams, which imparted higher cognitive load, adding an additional processing constraint on the unattended information. Despite this, both stimulus types were still processed somewhat deeply, leading to above chance recognition rates, overall, during the recognition test and no difference between stimulus types. Thus, it is likely that both auditory words and sounds used in the current experiment were treated in a similar manner by the cognitive system when unattended and presented in the auditory modality.

8. Experiment 2: Cross-Modal Conditions

Experiment 1 focused on exploring potential differences in the facilitated processing for ignored lexical and non-lexical information under unimodal conditions only (i.e., visual or auditory presentations). However, we live in a multisensory environment and are rarely presented with information to just one sensory modality at a time. Therefore, investigating how our cognitive system processes lexical and non-lexical information under multisensory conditions provides a more comprehensive and ecological understanding of information processing in a dynamic environment, in addition to providing information as to how the sensory modalities interact with one another when attending to and perceiving information in our environment. Therefore, Experiment 2 examined the extent to which processing for ignored lexical and non-lexical information might be facilitated when initially presented under multisensory (i.e., cross-modal) conditions.

8.1. Experiment 2a: Cross-Modal Ignored Visual Conditions

Experiment 2a examined the extent to which processing for ignored visual lexical and non-lexical information might be facilitated when initially presented with an attended auditory stream. Here, separate groups of participants were presented with a compound auditory/visual stream containing both lexical (i.e., auditory words or written words) and non-lexical items (i.e., sounds or pictures). They were asked to monitor the auditory dimension (i.e., sounds or auditory words) for task-relevant targets while ignoring the visual stream (i.e., written words or pictures). After completing this attention-demanding primary task, the extent to which the ignored auditory items may have been processed were evaluated via a surprise recognition test, administered in the auditory modality (see stimuli and procedure sections for details).

8.1.1. Participants

A total of 86 naïve, English speaking, young adults were recruited for Experiment 2a in the same manner as that of Experiment 1 (see section 7.1.1 for details). Of the 86 participants recruited, 40 (n = 40, *M* age = 19.0, 27 female) were randomly assigned to ignore non-lexical visual information (i.e., attend auditory words / ignore pictures condition) while the remaining 46 (n = 46, *M* age = 19.5, 33 female) were assigned to ignore lexical visual information (i.e., attend sounds / ignore written words condition).

Of the 40 assigned to the attend auditory words / ignore pictures condition, six participants were removed from data analysis due to having a high miss rate (i.e., 1*SD* above the mean) and one participant was removed for providing uniform responses (i.e., all 'no') during the surprise recognition test. Of the 46 assigned to the attend sounds / ignore written words condition, seven participants were removed from data analysis due to having a high miss rate during the primary task, one was removed for failing to complete the primary task, and two were removed for providing uniform responses during the surprise recognition test.

The reported analyses include the remaining 33 participants in the attend auditory words / ignore pictures condition (n = 33, *M* age = 18.97, 21 female) and the remaining 36 participants in the attend sounds / ignore written words condition (n = 36, *M* age = 19.44, 27 female). All participants were presented with informed consent (see appendix B) prior to beginning the experiment and debriefed upon completion (see appendix C).

8.1.2. Stimuli

Visual and auditory stimuli were identical to those presented in Experiments 1a and 1b but presented under cross-modal conditions rather than unimodal (see sections 7.1.2 and 7.2.2 for details).

8.1.2.1. Attended Sounds / Ignored Written Words

The same 50 auditory sounds presented in Experiment 1b (attended sounds / ignored auditory words, see section 7.2.2.2. for details) were presented in a compound auditory/visual rapid serial presentation with the eight selected written words presented in Experiment 1a (attended pictures / ignored written words, see sections 7.1.2.2. for details). All written words were presented an equal number of times and eight versions of the experiment were created in which each of the eight written words served as the TA item while the remaining seven written words served as the NA items.

8.1.2.2. Attended Auditory Words / Ignored Pictures

The same 50 auditory words presented in Experiment 1b (attended auditory words / ignored sounds, see section 7.2.2.1. for details) were presented in a compound auditory/visual rapid serial presentation with the eight selected pictures presented in Experiment 1a (attended written words / ignored pictures, see sections 7.1.2.1. for details). All pictures were presented an equal number of times and eight versions of the experiment were created in which each of the eight pictures served as the TA item while the remaining seven pictures served as the NA items.

8.1.2.3. Surprise Recognition Test

Recognition test items were presented in the visual modality only. As such, presentation parameters for previously ignored items (pictures or written words) were identical to those of Experiment 1a (see section 7.1.2.3. for details).

8.1.3. Procedure

As in Experiment 1b, participants were asked to attend to either the sounds or auditory words presented in the auditory stream and respond to immediate target repetitions by pressing the left mouse button with their preferred hand. However, in Experiment 2a, participants were asked to ignore the simultaneously presented visual stream (pictures or written words depending on condition). As before, each compound stimulus was presented for 350ms with a 150ms ISI (i.e., blank screen and silence) for a stimulus onset asynchrony (SOA) of 500ms (i.e., identical to Experiments 1a and 1b, see Figure 19). Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task (verified by experimenter observation and verbal confirmation of the participant). The surprise recognition test was administered to all participants immediately after the primary task in the same way that it was in Experiments 1a and 1b.

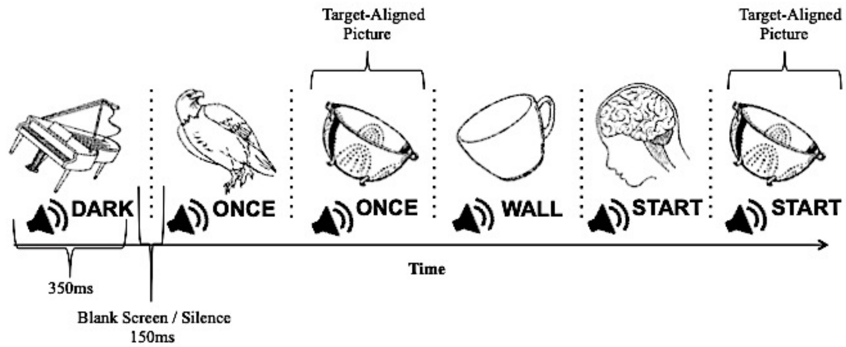


Figure 19. Schematic representation of the primary task, presented cross-modally, in the attended auditory words and ignored pictures condition. Auditory words were presented binaurally over headphones while pictures were simultaneously presented on the screen. In the schematic, auditory words are represented in bold, capitalized letters, along the bottom row. Immediate repetitions in the auditory word stream served as the target in the attended auditory dimension (e.g., “once” and “start”) while the pictures were the ignored items. Pictures appearing with immediate auditory word repetitions were the TA items (e.g., the colander) all other pictures were NA items. Notice that the TA item was always the same. However, all ignored items were presented an equal number of times during the entirety of the primary task.

8.1.4. Predictions

Regarding the primary task, as with previous experiments, above chance performance was expected for both groups. Furthermore, significant differences in d’ rates between groups (i.e., auditory words vs. sounds) were not expected due to the highly salient, easy to identify, and familiar stimuli used in the primary task. Therefore, accuracy was expected to be relatively high – meaning participants would have little difficulty identifying and responding to auditory target repetitions in the rapid serial compound auditory/visual stream within the allotted time frame (i.e., 1000ms after stimulus onset). However, it was still anticipated that participants would respond faster to sound repetitions compared to word repetitions during the primary task.

Considering the surprise recognition test, as with Experiments 1a and 1b, a main effect for temporal alignment was, again, expected. That is, temporally aligning an ignored visual item (i.e., picture or written word) with an attended task-relevant auditory target was predicted to result in facilitated processing for the ignored, task-irrelevant, information. This facilitated processing would be reflected in the surprise recognition test, with participants demonstrating overall higher recognition rates for TA items compared to NA items, regardless of stimulus type.

Furthermore, it is still assumed that, even under cross-modal conditions, ignored pictures will be processed more extensively compared to ignored written words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980; Tipper, 1985; Tipper & Driver, 1988). Thus, a main effect of stimulus type was predicted, meaning that recognition rates for previously ignored pictures would be significantly higher compared to previously ignored written words. However, it should be noted that many studies have demonstrated that cross-modal presentations improve performance across a variety of attention-demanding tasks (Busse et al., 2005, Kim, Seitz, & Shams, 2008; Laurienti, Burdette, Maldjian, & Wallace, 2006; Shams & Seitz, 2008; Spence & Driver, 1996, 1997; Spence et al., 2001; Sinnett et al., 2006; Van der Burg et al., 2008; Van der Burg et al., 2011). Therefore, due to the cross-modal nature of stimulus presentation, it was also possible that participants would exhibit overall high performance during the surprise recognition test across stimulus types leading to potential ceiling effects and similar response rates between groups.

Finally, it was difficult to predict how stimulus type and sensory modality would interact under these conditions. One possible explanation for the observed facilitatory effects of cross-modal stimulus presentation in other research (Busse et al., 2005; Kim et al., 2008; Laurienti et al., 2006; Shams & Seitz, 2008; Spence & Driver, 1996, 1997; Spence et al., 2001; Sinnett et al., 2006; Van der Burg et al., 2008; Van der Burg et al., 2011) has centered around the idea of segregated, or at least partially segregated, attentional resources across sensory modalities (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco, Morein-Zamir, & Kingstone, 2005; Wickens, 1984). This potential segregation of attentional resources would result in the auditory and visual information being processed separately (arguably in parallel), thereby allowing our cognitive system to successfully process more information under cross-modal circumstances, compared to unimodal presentations where a single reservoir of attentional resources would theoretically deplete more quickly. As a result, participants would be able to successfully monitor the auditory stream for task-relevant targets and still have enough attentional resources available to process the visual information, despite differences in stimulus type and having been instructed to explicitly ignore this information.

Indeed, cross-modally enhanced processing for unattended information has been observed. For example, Sinnett and colleagues (2006) found reduced levels of inattentional blindness (and deafness) for previously ignored items under cross-modal conditions using nearly identical stimuli and presentation conditions to those presented here. These findings were interpreted as evidence of attentional resources that are partially segregated across sensory modalities. As a result, ignored information was processed to a greater extent under cross-modal conditions, leading to higher recognition rates for previously ignored items, compared to recognition rates for ignored information presented in only unimodal conditions. Thus, in addition to temporal-alignment with a task-relevant target, cross-modal presentation of information during the primary task would further enhance the extent to which unattended information may be processed.

8.1.5. Results

8.1.5.1. Primary Task

Performance on the primary task was evaluated in the same manner as Experiment 1 (see section 7.1.5.1. for details). Analyses of primary task accuracy revealed that participants in both groups obtained a proportion of hits significantly above chance (12%) [attend auditory words / ignored pictures (i.e., attend auditory words): $M = 0.43$, $SE = 0.03$, $t(32) = 10.68$, $p < 0.001$; attend sounds / ignored written words (i.e., attend sounds): $M = 0.61$, $SE = 0.02$, $t(35) = 23.69$, $p < 0.001$].

Participants in the attend auditory words condition had significantly lower d' scores compared to those in the attend sounds condition [attend auditory words: $M = 3.11$, $SE = 0.11$ vs. attend sounds: $M = 3.49$, $SE = 0.05$, $t(72) = 5.51$, $p < 0.001$] (see Figure 20), suggesting a lower level of target detection sensitivity among those in the attend auditory words condition compared to those in the attend sounds condition. All analyses met Bonferroni corrected significance levels ($p < 0.01$)[11].

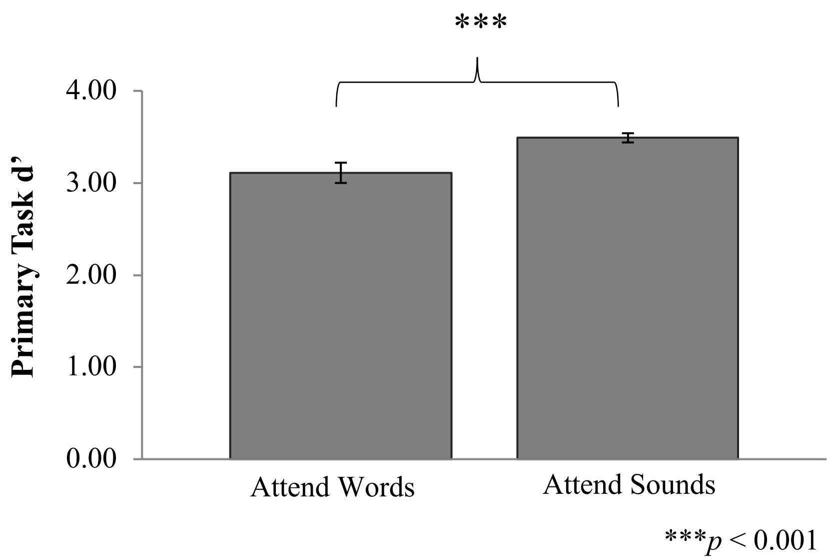


Figure 20. d' rates for target identification during the primary task in Experiment 2a. The primary task was presented in the auditory modality, while ignored items were presented in the visual modality. "Attend Words" indicates the condition in which participants monitored the RSAP stream for immediate word repetitions while ignoring visually presented pictures and "Attend Sounds" indicates the condition in which participants monitored the RSAP stream for immediate sound repetitions while ignoring visually presented words. Error bars represent the standard error for each variable. Those in the attend words condition had a d' score significantly lower ($M = 3.11$, $SE = 0.11$) than those in the attend sounds condition ($M = 3.49$, $SE = 0.05$, $p <$ 0.001), suggesting lower sensitivity to target identification among those in the attend words condition.

[11] Bonferroni corrections for primary task performance included four main analyses (0.05/4 = 0.01): two t-tests on hit rates against chance, one between group t-test on d' rates, and one between group t-test on RTs.

Finally, in order to evaluate processing speed for each stimulus type (i.e. pictures or written words) presented during the primary task, participants' RTs to identified targets were aggregated and compared between groups. Interestingly, those in the attend auditory words condition responded to targets significantly slower compared to those in the attend sounds condition [attend written words: M = 441ms, SE = 3.49 vs. attend sounds: M = 403ms, SE = 5.73,.$t(67) = 5.75$, $p < 0.001$], which met Bonferroni corrected significance levels ($p < 0.01$) (see Figure 21).

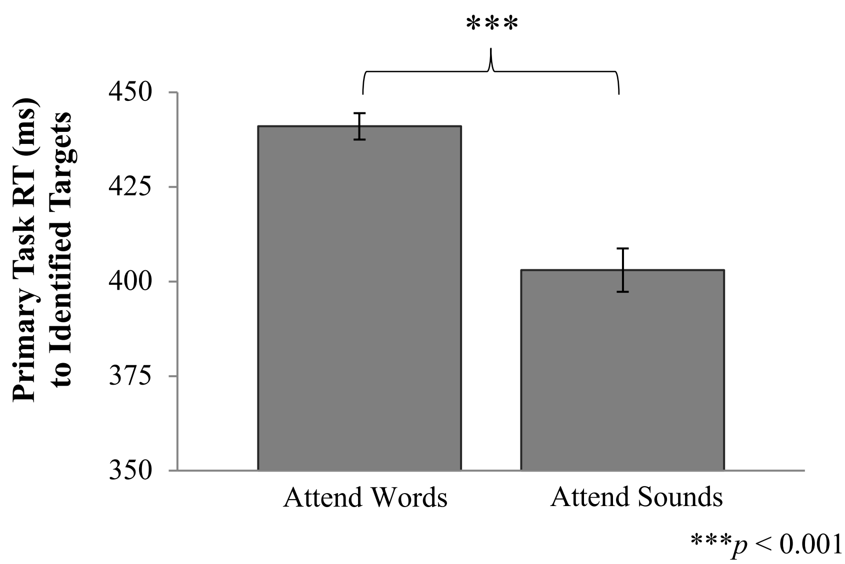


Figure 21. RT (ms) to identified auditory targets during the primary task in Experiment 2a. "Attend Words" indicates the condition in which participants monitored the RSAP stream for immediate word repetitions while ignoring visually presented pictures and "Attend Sounds" indicates the condition in which participants monitored the RSAP stream for immediate sound repetitions while ignoring visually presented words. Error bars represent the standard error for each variable. Participants in the attend words condition responded to targets significantly slower (M = 441ms, SE = 3.49) compared to those in the attend sounds condition (M = 403ms, SE = 5.73, $p < 0.001$).

8.1.5.2. Surprise Recognition Test

Performance on the surprise recognition was evaluated in the same manner as Experiment 1a (see section 7.1.5.2. for details). Analyses of surprise recognition test accuracy revealed that participants in both groups obtained an accuracy score (Hits + CR) significantly above chance (50%) [attend sounds / ignored written words (i.e., ignored written words): $M = 0.88$, $SE = 0.02$, $t(35) = 22.18$, $p < 0.001$; attend auditory words / ignored pictures (i.e., ignored pictures): $M = 0.95$, $SE = 0.02$, $t(32) = 25.32$, $p < 0.001$] (see Figure 22). Both variables failed to meet the assumption of normality; therefore, these analyses were corroborated by One-Sample Wilcoxon Signed-Ranks Tests [ignored written words: Z = 5.27, $p < 0.001$; ignored pictures: Z = 5.11, $p < 0.001$]. All tests met Bonferroni corrections for multiple comparisons ($p < 0.02$)[12].

[12] Bonferroni corrections for overall accuracy performance during the surprise recognition test included three main analyses (0.05/3 = 0.02): two t-tests on accuracy scores (hits + CRs) against chance and one between-group t-test on d' rates.

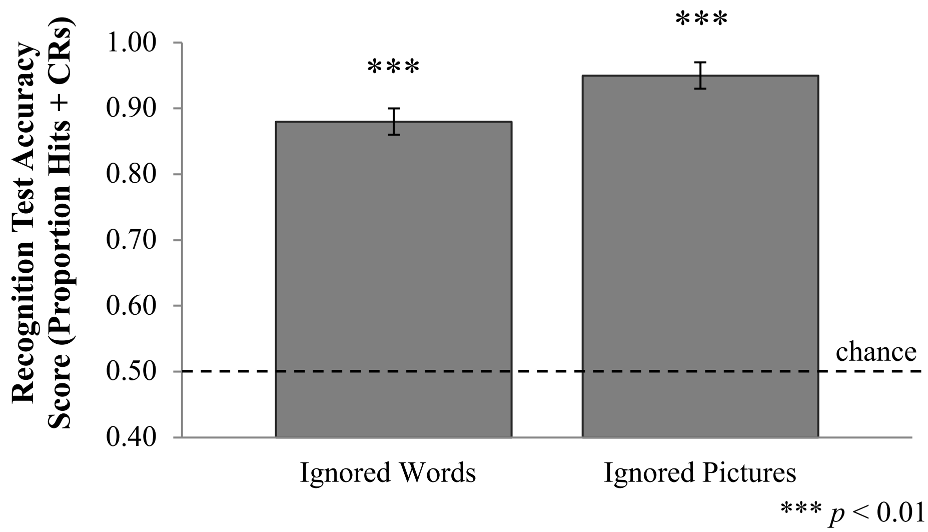


Figure 22. Accuracy scores (i.e., proportion of hits + CRs) for the surprise recognition test in Experiment 2a. The primary task was presented in the auditory modality, while ignored items were presented in the visual modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored written words and "Ignored Pictures" indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. Both groups had an accuracy score significantly above chance (i.e., 0.50, $p < 0.001$), suggesting successful completion of the surprise recognition test.

Those in the ignored written words condition had a d' significantly lower compared to those in the ignored pictures condition, according to conventional standards ($p < 0.05$) [ignored written words: $M = 4.71$, $SE = 0.35$ vs. ignored pictures: $M = 5.85$, $SE = 0.32$, $t(67) = 2.39$, $p = 0.02$], but not when corrected for multiple comparisons ($p < 0.02$). The d' variables for both groups failed to meet the assumption of normality; therefore, this analysis was corroborated by Mann-Whitney U Test [U = 3.17, $p = 0.002$], which did meet corrected levels of significance levels ($p < 0.02$). Taken together, these results suggest a lower level of target detection sensitivity among those in the ignored words condition compared to those in the ignored pictures condition (see Figure 23).

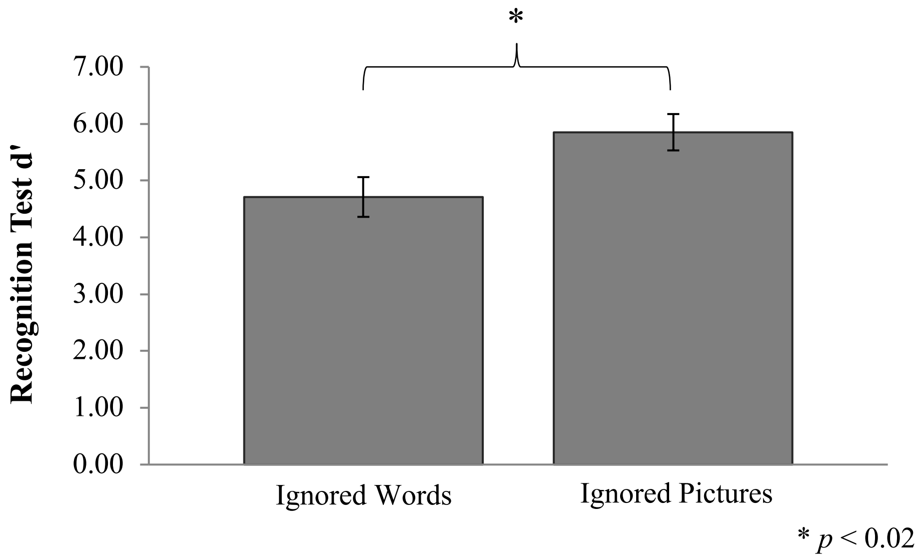


Figure 23. d’ rates for target identification during the surprise recognition test in Experiment 2a. The primary task was presented in the auditory modality, while ignored items were presented in the visual modality. “Ignored Words” indicates the condition in which participants were tested on previously ignored written words and “Ignored Pictures” indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. Those in the ignored words condition had a d’ score significantly lower (M = 4.71, SE = 0.35) than those in the ignored pictures condition (M = 5.85, SE = 0.32, $p <$ 0.02), suggesting lower sensitivity to target identification among those in the ignored words condition during the surprise recognition test.

Regarding the critical analysis, a two-factor (2x2) ANOVA was conducted on recognition performance for the surprise recognition test, with focus of attention (ignored written words or ignored pictures) as the between subjects factor and target-alignment (TA or NA) as the within subjects factor. There was a main effect for target-alignment indicating that, overall, TA items [M = 0.96, SE = 0.03] were recognized significantly more often than NA items [$M = 0.85$, $SE = 0.02$, $F(1,67) = 11.94$, $p < 0.001$]. There was a main effect for group type indicating that ignored written words [$M = 0.79$, $SE = 0.03$] were recognized significantly less often than ignored pictures [$M = 0.95$, $SE = 0.01$, $F(1,67) = 15.16$, $p < 0.001$]. Finally, there was no interaction [$F(1,67) = 1.95$, $p = 0.167$] (see Figure 24). Despite a lack of interaction, planned comparisons described below explored recognition rates for TA and NA items specifically within, and between, each condition (ignored written words or ignored pictures).

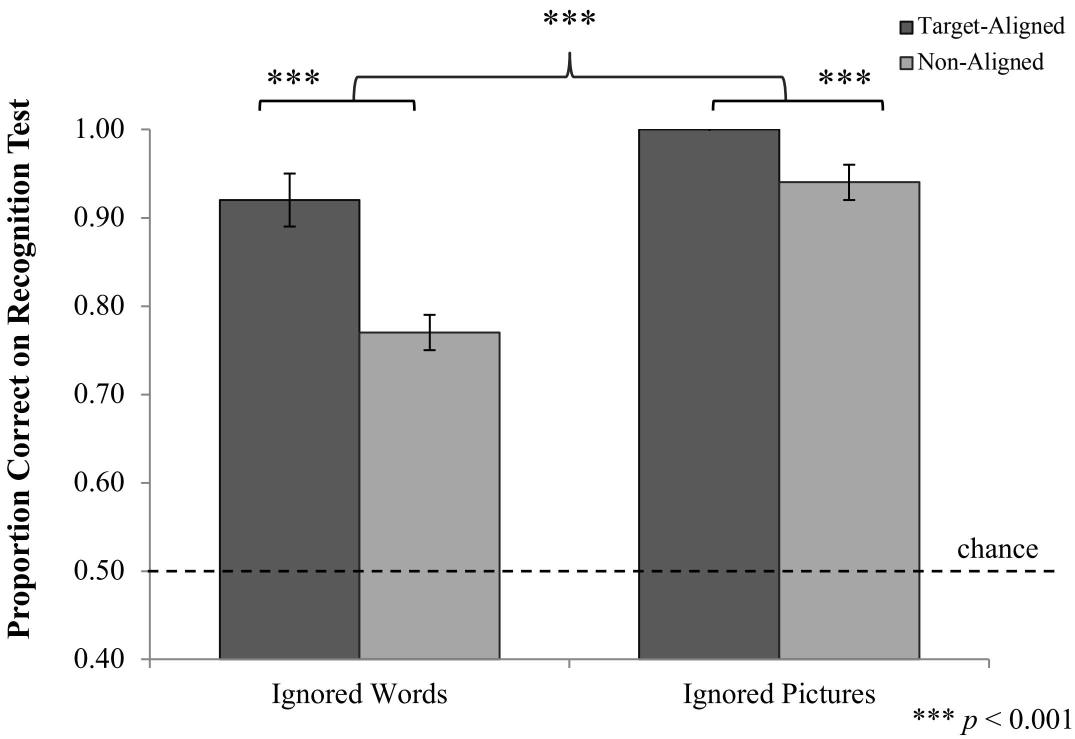


Figure 24. Main effects from the critical analysis of the surprise recognition test in Experiment 2a. The primary task was presented in the auditory modality, while ignored items were presented in the visual modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored written words and "Ignored Pictures" indicates the condition in which participants were tested on previously ignored pictures. Error bars represent the standard error for each variable. Note that TA pictures had a mean of one and a standard deviation of zero ($M = 1.00$, $SD = 0.00$). All items were recognized at rates significantly above chance (0.50 – significance not shown). There was a main effect for target-alignment ($p < 0.001$), suggesting that TA items were recognized more often than NA items. There was also a main effect for stimulus type ($p < 0.001$), suggesting that previously ignored pictures were recognized more often that previously ignored written words. There was no interaction.

8.1.5.2.1. Ignored Written Words

Overall, participants were able to recognize the previously ignored written words (TA and NA) statistically better than chance (50%) [$M = 0.79$, $SE = 0.03$, $t(35) = 9.90$, $p < 0.001$]. Recognition for TA words [$M = 0.92$, $SE = 0.05$, $t(35) = 8.92$, $p < 0.001$], and NA words [$M = 0.77$, $SE = 0.03$, $t(35) = 8.46$, $p < 0.001$] was each better than chance. Both variables failed to meet assumptions of normality; therefore, the TA word analysis was corroborated by a Binomial Test [the observed proportion of TA words (0.92) was significantly higher than the expected proportion (0.50), $p < 0.001$] and the NA word analysis was corroborated by a One-Sample Wilcoxon Signed-Ranks Test [NA words: Z = 4.96, $p < 0.001$]. All tests met Bonferroni corrections for multiple comparisons ($p < 0.005$)[13]. Finally, when compared to each other, TA words were recognized significantly more often than NA words according to conventional standards ($p < 0.05$) [$t(35) = 2.63$, $p = 0.01$], corroborated by Wilcoxon Signed Ranks Test [Z = 2.51, $p = 0.01$], though this did not meet corrected significance levels ($p < 0.005$).

8.1.5.2.2. Ignored Pictures

Overall, participants were able to recognize the previously ignored pictures (TA and NA) statistically better than chance (50%) [$M = 0.95$, $SE = 0.01$, $t(32) = 30.99$, $p < 0.001$], which met Bonferroni corrections for multiple comparisons ($p < 0.005$). A One-Sample *t*-test against chance on recognition rates for TA pictures could not be performed because the variable had a mean of one and standard error of zero ($M = 1.0$, $SE = 0.00$), indicating a possible ceiling effect; therefore, the TA picture recognition was evaluated by a Binomial Test only [the observed proportion of TA pictures (1.00) was significantly different than the expected proportion (0.50), $p < 0.001$]. Recognition for NA pictures [$M = 0.94$, $SE = 0.02$, $t(32) = 26.66$, $p < 0.001$] was significantly better than chance and met the Bonferroni corrected significant level ($p < 0.005$). NA pictures failed to meet assumptions of normality; therefore, this analysis was corroborated by a One-Sample Wilcoxon Signed-Rank Test [NA sounds: Z = 5.20, $p < 0.001$]. Finally, when compared to each other, TA pictures were recognized

[13] Bonferroni corrections for post hoc investigations included ten main analyses (0.05/10 = 0.005): four t-tests on TA and NA items for the ignored words condition, four t-test on TA and NA items for the ignored pictures condition, and two t-test on TA and NA items between conditions.

significantly more often than NA pictures [$t(32) = 3.68$, $p < 0.001$], corroborated by Wilcoxon Signed Ranks Test [$Z = 3.36$, $p = 0.001$], both of which met Bonferroni corrected significance levels ($p < 0.005$).

8.1.5.2.3. Analysis by Target-Alignment

Comparing performance on the surprise recognition test for TA and NA items between each group revealed that there was no significant difference in recognition rates for TA items between the ignored written words condition [$M = 0.92$, $SE = 0.05$] and ignored pictures condition [$M = 1.0$, $SE = 0.00$, $t(67) = 1.78$, $p = 0.08$], corroborated by Mann-Whitney U Test [$U = 1.68$, $p = 0.09$]. However, NA items for ignored written words [$M = 0.77$, $SE = 0.03$] were recognized significantly less often than NA items for ignored pictures [$M = 0.94$, $SE = 0.02$, $t(67) = 21.70$, $p < 0.001$], corroborated by Mann-Whitney U Test [$U = 377$, $p < 0.001$] (see Figure 25), both of which met the Bonferroni corrected significance level ($p < 0.005$).

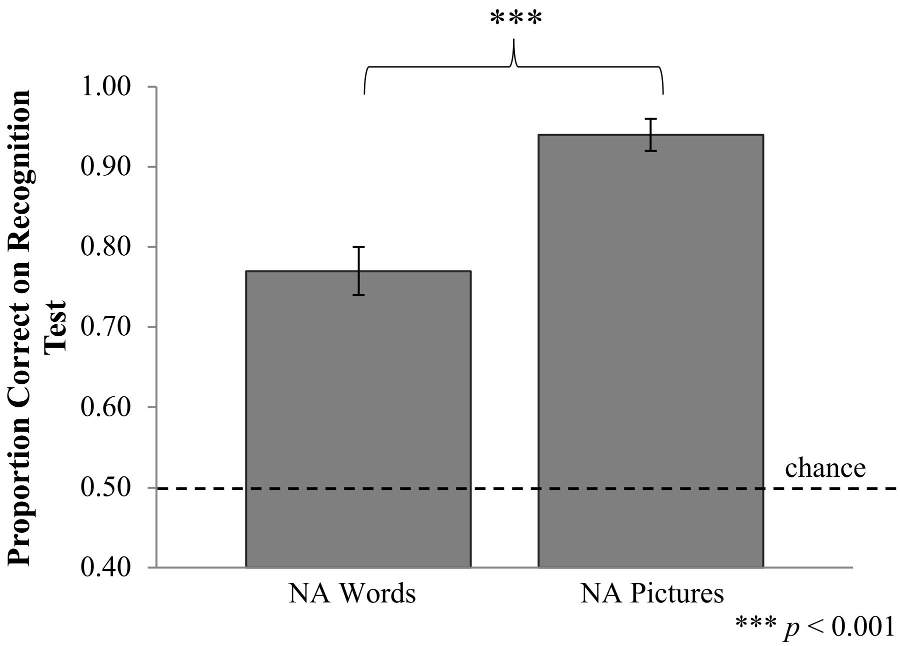


Figure 25. Recognition rates for NA items during the surprise recognition test in Experiment 2a. The primary task was presented in the auditory modality, while ignored items were presented in the visual modality. “NA Words” indicates recognition rates for non-aligned written words from the “Ignored Words” condition and “NA Pictures” indicates recognition rates for non-aligned pictures from the “Ignored Pictures” condition. Error bars represent the standard error for each variable. Both NA items were recognized at rates significantly above chance (0.50 – significance not shown). Those in the ignored words condition recognized significantly fewer NA items (M = 0.62, SE = 0.03) than those in the ignored pictures condition (M = 0.81, SE = 0.03, p < 0.001), suggesting that NA pictures were facilitated to a greater extent compared to NA words.

8.1.6. Discussion: Experiment 2a

Experiment 2a evaluated potential differences in the facilitated processing of previously unattended visual lexical (i.e., written words) and non-lexical (i.e., pictures) stimuli under cross-modal auditory/visual conditions in an attention-demanding task. During the primary task, participants were asked to attend only to the auditory stream while ignoring the simultaneously presented visual stream. As predicted, both groups obtained a proportion of hits significantly above chance, which suggests that participants in both conditions were able to successfully identify auditory targets in the cross-modal auditory/visual stream. Those in the attend sounds condition had significantly higher d' scores compared to those in the attend auditory words condition. This finding suggests greater sensitivity for signal detection among those in the attend sounds condition compared to those in the attend auditory words condition. Interestingly, those in the attend sounds condition also responded to targets significantly faster compared to those in the attend auditory words condition. Therefore, when participants were monitoring the auditory stream for immediate repetitions while ignoring a simultaneously presented visual stream, sound targets appear to be identified, and responded to, more accurately compared to auditory word targets.

The results from the primary repetition detection task suggest that presenting the two streams under cross-modal conditions may have lessened the level of cognitive load (Cartwright-Finch & Lavie, 2007; Lavie, 2005, 2010) associated with the task by allowing participants to draw on partially segregated attentional resources across sensory modalities (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984). That is, by presenting the unattended stream to the visual modality, auditory attentional resources were free to process the sound or auditory word stream in a semi-separate manner. This additional processing capacity may have altered the way in which the cognitive system handles the sound stream compared to the auditory word stream, such that attended sounds were, indeed, processed more readily, leading to higher accuracy and faster RTs, but only under cross-modal conditions.

Regarding the surprise recognition test for previously unattended visual items (i.e., ignored written words or ignored pictures), both groups obtained accuracy scores significantly above chance, suggesting that participants in both groups were able to distinguish between previously seen items and foil items during the recognition test. However, those in the ignored pictures condition had an accuracy score significantly higher compared to those in the ignored written words condition, as well as a higher d' (but note, the d' comparison failed to meet the criterion for multiple comparisons). These findings suggest that those in the ignored pictures condition were better able to distinguish between previously seen items and foil items compared to those in the ignored words condition. Such a finding aligns well with the idea that processing for pictures is more likely to be facilitated, even when ignored, and when presented under cross-modal conditions, leading to higher recognition rates – in general – during the surprise recognition test compared to words.

Considering the critical analysis on the surprise recognition test data, a two-factor (2x2) repeated measures ANOVA was conducted on participant responses with target-alignment (TA vs. NA) as the within subjects factor and stimulus type (ignored pictures vs. ignored written words) as the between subjects factor. There was a main effect for target-alignment suggesting that, overall, TA items (words and pictures) were recognized significantly more often than NA items (words and pictures) during the surprise recognition test. Thus, it appears that when presented under cross-modal conditions, temporal alignment between an ignored visual item and an attended auditory target (i.e., sound or auditory word) leads to facilitated processing of the ignored information in the visual modality, resulting in higher recognition rates when that information is encountered again at a later time, compared to equally presented, but unattended items that were not paired with an attended auditory target (i.e., NA items), regardless of stimulus type (words or pictures).

Next, the ANOVA revealed a main effect for stimulus type, indicating that previously ignored pictures (TA and NA) were recognized significantly more often than previously ignored words. Presenting the primary task under cross-modal conditions was expected to free up additional attentional resources for attended and ignored items (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984), leading to higher recognition rates for both stimulus types, overall, during the surprise recognition test. However, the fact that pictures are still recognized at significantly higher rates compared to words, lends further credence to the notion that pictures are processed more readily compared to words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980), even when unattended (Tipper, 1985, see also Tipper & Driver 1988) and when initially presented under cross-modal conditions.

Considering the question of whether recognition rates for TA and NA items would differ depending on stimulus modality (i.e., written words vs. pictures), predictions were, again, tenuous. As mentioned before, the cross-modal nature of the primary task was expected to lead to enhanced processing capacity within each sensory modality (i.e., vision and audition, see Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984). This alone was expected to result in improved recognition rates during the surprise recognition test compared to unimodal presentations. Furthermore, based on previous work demonstrating the robust effect that target-alignment has on enhancing the recognition of previously ignored items in the visual modality (Dewald et al., 2013, Seitz & Watanabe, 2003; Walker et al., 2014; Walker et al., 2017; Watanabe et al., 2001), and findings indicating that pictures may be processed more readily compared to words (Amit et al., 2009; Carr, McCauley, Sperber, & Parmelee, 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980), it seemed plausible that an interaction would occur, however no such interaction was observed. Thus, it appears that even under cross-modal conditions, the facilitatory effects of target-alignment remain constant despite noted differences regarding how each stimulus type (words and pictures) are processed by the human cognitive system.

Finally, despite the fact that an interaction was not observed, pre-planned analyses investigated the relationship between recognition rates for TA and NA items within and between each stimulus condition. Participants who were tested on previously ignored written words recognized both TA and NA items at rates significantly above chance. TA words were recognized significantly more often than NA words under conventional levels of significance but not when corrected for multiple comparisons. Next, participants who were tested on previously ignored pictures recognized both TA and NA items at rates significantly above chance, and TA pictures were recognized significantly more often than NA pictures. Finally, comparing performance on recognition rates for TA and NA items *between* stimulus types revealed no significant difference in recognition for TA items between pictures and words, however, NA pictures were recognized significantly more often than NA words.

Combined, the findings from Experiment 2a further suggest facilitated processing for pictures compared to words when actively ignored and, again, support the notion that facilitation by target-alignment for unattended items proceeds in a similar manner between lexical and non-lexical stimuli in the visual modality. This trend appears to persist even when the primary task was presented under cross-modal conditions. Thus, even when additional resources were available to process the unattended visual items (pictures and words) while attending to the auditory stream, pictures still appear to be processed more extensively and subsequently recognized at higher rates compared to words. Furthermore, temporal pairing with an attended task target appears to further facilitate processing for unattended TA items compared to NA items, regardless of stimulus type, under cross-modal conditions.

8.2. Experiment 2b: Cross-Modal Ignored Auditory Conditions

The final experiment of this doctoral dissertation focused on exploring the extent to which processing for explicitly ignored lexical and non-lexical auditory information may be facilitated when attention is directed toward a visual stream. As with Experiment 2a, participants were presented with a compound auditory/visual stream containing both lexical (i.e., auditory or written words) and non-lexical (i.e., sounds or pictures) items. Here, they were asked to monitor the visual dimension (i.e., pictures or written words) for task-relevant targets while ignoring the auditory stream (i.e., auditory words or sounds). After completing this attention-demanding primary task, the extent to which the ignored auditory items had been processed were evaluated via a surprise recognition test (see stimuli and procedure for details).

8.2.1. Participants

A total of 99 naïve, English speaking, young adults were recruited for Experiment 2b in the same manner as that of Experiment 1 (see section 7.1.1. for details) and were able to utilize devices designed to correct visual or auditory impairments. Of the 99 participants recruited, 46 (n = 46, *M* age = 20.5, 29 female) were randomly assigned to ignore non-lexical auditory information (i.e., attend written words / ignore sounds) while the remaining 53 (n = 53, *M* age = 20.5, 30 female) were assigned to ignore lexical auditory information (i.e., attend pictures / ignore auditory words).

Of the 46 assigned to the attend written words / ignore sounds condition, nine participants were removed from data analysis due to having a high miss rate (i.e., 1*SD* above the mean) and one participant was removed due to having a high FA rate (i.e., 1*SD* above the mean) during the primary task. As with previous experiments, the stringent performance criteria were applied to ensure participants' attention was fully focused on the primary task. Of the 53 assigned to the attend pictures / ignore auditory words condition, eight participants were removed from data analysis due to having a high miss rate during the primary task, and one was removed for providing uniform responses (i.e., all 'no') during the surprise recognition test. The reported analysis includes the remaining 36 participants in the attend written words / ignore sounds condition (n = 36, *M* age = 20.6, 23 female) and the remaining 44 participants in the attend pictures / ignore auditory words condition (n = 44, *M* age = 19.9, 23 female). All participants were presented with informed consent (see appendix B) prior to beginning the experiment and debriefed upon completion (see appendix C).

8.2.2. Stimuli

Visual and auditory stimuli were identical to those presented in Experiments 2a but with the opposite attended and ignored dimensions. See section 8.1.2. for details.

8.2.2.1. Attended Pictures / Ignored Auditory Words

The same 50 pictures presented in Experiment 1a (attended pictures / ignored written words, see section 7.1.2.2. for details) were presented in a compound auditory/visual rapid serial presentation with the eight selected auditory words in Experiment 1b (attended sounds / ignored auditory words, see section 7.2.2.2. for details). All auditory words were presented an equal number of times and eight versions of the experiment were created in which each of the eight auditory words served as the TA item while the remaining seven auditory words served as the NA items.

8.2.2.2. Attended Written Words / Ignored Sounds

The same 50 written words presented in Experiment 1a (attended written words / ignored pictures, see section 7.1.2.1. for details) were presented in a compound auditory/visual rapid serial presentation with the eight selected sounds in Experiment 1b (attended auditory words / ignored sounds, see section 7.2.2.1. for details). All sounds were presented an equal number of times and eight versions of the experiment were created in which each of the eight sounds serves as the TA item while the remaining seven sounds serve as the NA items.

8.2.2.3. Surprise Recognition Test

Recognition test items were presented in the auditory modality only. As such presentation parameters for previously ignored items (sounds or auditory words) were identical to those of Experiment 1b (see section 7.2.2.3. for details).

8.2.3. Procedure

Participants were asked to attend to either the pictures or the written words presented in the visual stream and respond to immediate target repetitions by pressing the left mouse button with their preferred hand while ignoring the simultaneously presented auditory stream (sounds or auditory words depending on condition). As before, each compound stimulus was presented for 350ms with a 150ms ISI (i.e., blank screen and silence) for a stimulus onset asynchrony (SOA) of 500ms (i.e., identical to Experiments 1a, 1b, and 2a, see Figure 26). Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task (verified by experimenter observation and verbal confirmation of the participant). As in all previous experiments, the surprise recognition test was administered to all participants immediately after the primary task in the exact same way, but this time entirely in the auditory modality (i.e., surprise tests for auditory words or sounds).

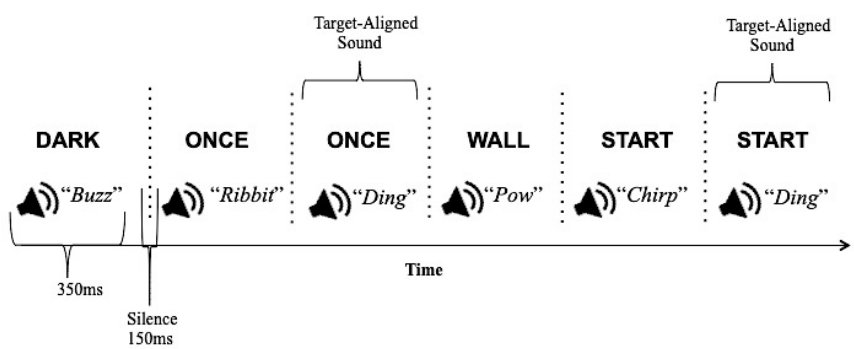


Figure 26. Schematic representation of the primary task, presented cross-modally, in the attended written words and ignored sounds condition. Written words were presented on the screen in bold, capitalized letters, while simultaneously presented sounds were presented binaurally over headphones. In the schematic, unattended sounds are represented in italicized letters and quotation marks along the bottom row. Immediate repetitions in the written word stream served as the target in the attended visual dimension (e.g., "once" and "start") while the sounds were the ignored items. Sounds appearing with immediate written word repetitions were the TA items (e.g., "*Ding"*), while all other sounds were NA items. Notice that the TA item was always the same. All ignored items were presented an equal number of times during the entirety of the primary task.

8.2.4. Predictions

As with the previous experiments reported in this doctoral dissertation, high performance on the primary task was certainly expected given that all items are highly salient and familiar. As such, participants were expected to exhibit similar performance rates across all primary task dimensions measured in Experiment 2b (i.e., hit rate, d', and RT).

In line with all previous experiments, a main effect of target-alignment was predicted. This finding would suggest that, even under the purportedly enhanced processing conditions of cross-modal stimulus presentation, the temporal alignment of an ignored auditory item with an attended task-relevant visual target leads to facilitated processing for the ignored information.

A main effect for stimulus type, in favor of non-lexical auditory information, was expected. This would suggest that ignored sounds are processed more extensively than ignored auditory words, leading to higher recognition rates during the surprise recognition test. This prediction was made based on previous research suggesting that the high frequency of selected words presented in these experiments may actually hinder processing for this type of information due to the activation of competing representational nodes. Conversely, high-frequency, sub-lexical sounds appear to enjoy facilitated processing because they do not appear to activate competitive representational nodes to the same extent (Vitevitch & Luce, 1998, 2016). However, due to the cross-modal presentation of primary task, allowing for a partial separation of attentional resources and subsequent facilitation of information processing as a result (Busse et al., 2005; Kim et al., 2008; Laurienti et al., 2006; Shams & Seitz, 2008; Spence & Driver, 1996, 1997; Spence et al., 2001; Sinnett et al., 2006; Van der Burg et al., 2008, 2011), participants may exhibit overall high performance on the surprise recognition test regardless of stimulus type, leading to ceiling effects.

Finally, as before, predicting the extent to which sensory modality and stimulus type may interact across the conditions presented in Experiment 2b was challenging. As before, cross-modal presentation of information during the primary task may further enhance the extent to which ignored information is processed and this may compound with target-alignment leading to additional facilitatory effects. Under these conditions, previously ignored TA sounds would be recognized at rates significantly above those of previously ignored TA auditory words. However, as noted earlier, the cross-modal presentation of items during the primary task may also lead to overall higher recognition rates across stimulus types leading to ceiling effects, meaning a significant interaction would not be observed.

8.2.5. Results

8.2.5.1. Primary Task

Performance on the primary task was evaluated in the same manner as Experiment 1 (see section 7.1.5.1. for details). Analyses of primary task accuracy revealed that participants in both groups obtained a proportion of hits significantly above chance (12%) [attend written words / ignored sounds (i.e., attend written words): M = 0.54, SE = 0.03, $t(35) = 14.31$, $p < 0.001$; attend pictures / ignored auditory words (i.e., attend pictures): M = 0.57, SE = 0.02, $t(43) = 23.84$, $p < 0.001$]. The proportion of hits for the attend written words variable failed to meet the assumption of normality; therefore, this analysis was corroborated by a One-Sample Wilcoxon Signed-Rank Test [$Z = 5.23$, $p < 0.001$]. All analyses met Bonferroni corrections for multiple comparisons ($p < 0.01$)[14]. There was no significant difference in d' rates between groups [attend written words: M = 3.66, SE = 0.12 vs. attend pictures: M = 3.83, SE = 0.11, $t(78) = 1.04$, $p = 0.301$], suggesting comparable levels of target detection sensitivity between groups (see Figure 27).

[14] Bonferroni corrections for primary task performance included four main analyses (0.05/4 = 0.01): two t-tests on hit rates against chance, one between group t-test on d' rates, and one between group t-test on RTs.

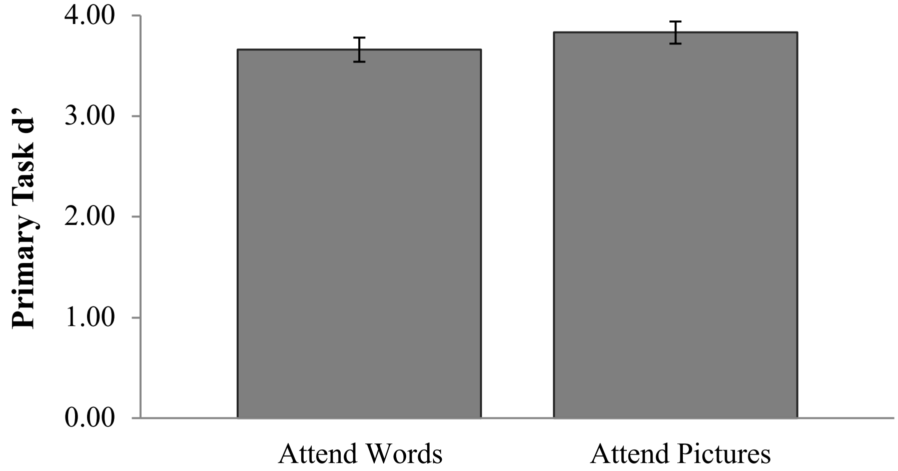


Figure 27. d' rates for target identification during the primary task in Experiment 2b. The primary task was presented in the visual modality, while ignored items were presented in the auditory modality. "Attend Words" indicates the condition in which participants monitored the RSVP stream for immediate word repetitions while ignoring aurally presented sounds and "Attend Pictures" indicates the condition in which participants monitored the RSVP stream for immediate picture repetitions while ignoring aurally presented words. Error bars represent the standard error for each variable. There was no significant difference in d' rates between groups, suggesting comparable levels of sensitivity to target identification between groups

Finally, in order to evaluate processing speed for each stimulus type (i.e. pictures or written words) presented during the primary task, participants' RTs to identified targets were aggregated and compared between groups. There was no significant difference in RT to identified targets during the primary task between groups [attend written words: $M = 423$, $SE = 4.04$ vs. attend pictures: $M = 422$, $SE = 2.49$, $t(78) = 0.183$, $p = 0.856$].

8.2.5.2. *Surprise Recognition Test*

Performance on the surprise recognition was evaluated in the same manner as all previous experiments (see section 7.1.5.2. for details). Analyses of surprise recognition test accuracy revealed that participants in both groups obtained an accuracy score (Hits + CR) significantly above chance (50%) [attend pictures / ignore auditory words (i.e., ignored auditory words): $M = 0.87$, $SE = 0.02$, $t(43) = 17.32$, $p < 0.001$, attend written words / ignore sounds (i.e., ignored sounds): $M = 0.92$, $SE = 0.01$, $t(35) = 29.54$, $p < 0.001$] (see Figure 28). The overall accuracy scores for both groups failed to meet the assumption of normality; therefore, these analyses were corroborated by One-Sample Wilcoxon Signed-Rank Tests [ignored auditory words: Z = 5.74, $p < 0.001$, ignored sounds: Z = 5.28, $p < 0.001$]. All tests met Bonferroni corrections for multiple comparisons ($p < 0.02$)[15].

[15] Bonferroni corrections for overall accuracy performance during the surprise recognition test included three main analyses (0.05/3 = 0.02): two t-tests on accuracy scores (hits + CRs) against chance and one between-group t-test on d' rates.

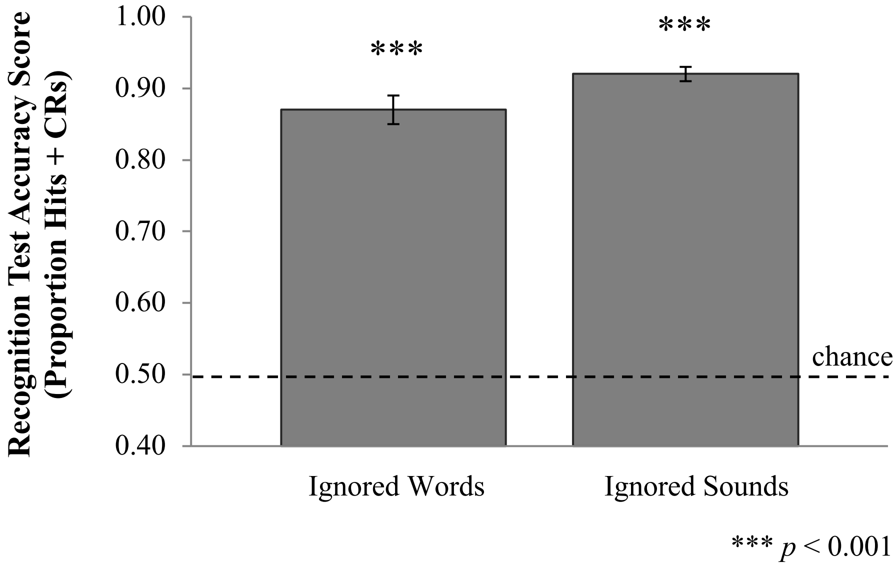


Figure 28. Accuracy scores (i.e., proportion of hits + CRs) for the surprise recognition test in Experiment 2b. The primary task was presented in the visual modality, while ignored items were presented in the auditory modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored auditory words and "Ignored Sounds" indicates the condition in which participants were tested on previously ignored sounds. Error bars represent the standard error for each variable. Both groups had an accuracy score significantly above chance (i.e., 0.50, $p < 0.001$), suggesting successful completion of the surprise recognition test.

There was no significant difference in d' between groups [ignored auditory words: $M = 4.95$, $SE = 0.38$ vs. ignored sounds: $M = 5.97$, $SE = 0.35$, $t(78) = 1.92$, $p = 0.06$]. Nonparametric analyses did reveal a significant difference by conventional standards ($p < 0.05$) [U = 2.23, $p = 0.03$], though this failed to meet the Bonferroni corrected significance level ($p < 0.02$). Taken together, these results suggest a comparable level of target detection sensitivity between those in the ignored auditory words condition and those in the ignored sounds condition.

Concerning the critical analysis, a two-factor (2x2) ANOVA was conducted on recognition performance for the surprise recognition test, with focus of attention (ignored auditory words or ignored sounds) as the between-subjects factor and target-alignment (TA or NA) as the within subjects factor. There was no main effect for target-alignment indicating that, overall, TA items [$M = 0.84$, $SE = 0.04$] were not recognized significantly more often than NA items [$M = 0.86$, $SE = 0.02$, $F(1,78) = 0.516$, $p = 0.475$]. There was no main effect for group type indicating that ignored auditory words [$M = 0.83$, $SE = 0.03$] were not recognized at rates significantly different from ignored sounds [$M = 0.87$, $SE = 0.03$, $F(1,78) = 0.597$, $p = 0.442$]. Finally, there was no interaction [$F(1,78) = 1.48$, $p = 0.226$] (see Figure 29). Despite a lack of interaction, planned comparisons described below explored recognition rates for TA and NA items specifically within, and between, each condition (ignored auditory words or ignored sounds).

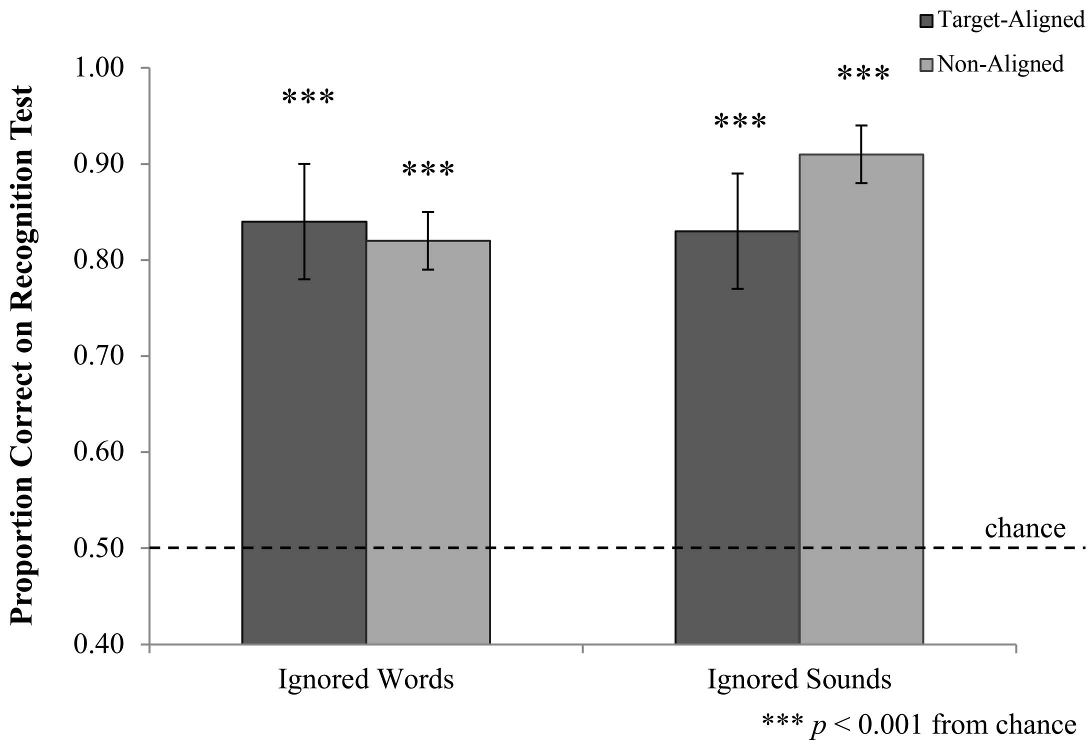


Figure 29. Results from the critical analysis of the surprise recognition test in Experiment 2b. The primary task was presented in the visual modality, while ignored items were presented in the auditory modality. "Ignored Words" indicates the condition in which participants were tested on previously ignored auditory words and "Ignored Sounds" indicates the condition in which participants were tested on previously ignored sounds. Error bars represent the standard error for each variable. There was no main effect for target-alignment, suggesting that TA items were not recognized more often than NA items. There was no main effect for stimulus type, suggesting that previously ignored sounds were not recognized more often that previously ignored auditory words. There was no interaction. All TA and NA items were recognized at rates significantly above chance (0.50, $p < 0.001$).

8.2.5.2.1. Ignored Auditory Words

Overall, participants were able to recognize the previously ignored auditory words (TA and NA) statistically better than chance (50%) [$M = 0.82$, $SE = 0.03$, $t(43) = 12.40$, $p < 0.001$]. Recognition for TA words [$M = 0.84$, *SE* = 0.06, *t*(43) = 6.11, $p < 0.001$], and NA words [*M* = 0.82, *SE* = 0.03, *t*(43) = 12.21, $p < 0.001$] was each better than chance. Both variables failed to meet assumptions of normality; therefore, the TA word analysis was corroborated by a Binomial Test [the observed proportion of TA words (0.84) was significantly higher than the expected proportion (0.50), $p < 0.001$] and the NA word analysis was corroborated by a One-Sample Wilcoxon Signed-Ranks Test [NA words: Z = 5.67, $p < 0.001$]. All tests met the Bonferroni corrected significance level ($p < 0.005$)[16]. Finally, when compared to each other, there was no significant difference in recognition rates between TA words and NA words [*t*(43) = 0.364, $p = 0.717$], corroborated by Wilcoxon Signed Ranks Test [Z = 1.11, $p = 0.266$].

8.2.5.2.2. Ignored Sounds

Overall, participants were able to recognize the previously ignored sounds (TA and NA) statistically better than chance (50%) [*M* = 0.90, *SE* = 0.03, *t*(35) = 15.78, $p < 0.001$]. Recognition for the TA sounds [*M* = 0.83, *SE* = 0.06, *t*(35) = 5.30, $p < 0.001$] and the NA sounds [*M* = 0.91, *SE* = 0.02, *t*(35) = 17. 34, $p < 0.001$] was each better than chance. Both variables failed to meet assumptions of normality; therefore, the TA sound analysis was corroborated by a Binomial Test [the observed proportion of TA sounds (0.83) was significantly higher than the expected proportion (0.50), $p < 0.001$] and the NA sound analysis was corroborated by a One-Sample Wilcoxon Signed-Ranks Test [NA sounds: Z = 4.96, $p < 0.001$]. All tests met the Bonferroni corrected significance level ($p < 0.005$). Finally, when compared to each other, there was no significant difference between recognition rates between TA sounds and NA sounds [*t*(35) = 1.34, $p = 0.186$], corroborated by Wilcoxon Signed Ranks Test [Z = 0.858, $p = 0.391$].

[16] Bonferroni corrections for post hoc investigations included ten main analyses (0.05/10 = 0.005): four t-tests on TA and NA items for the ignored words condition, four t-test on TA and NA items for the ignored pictures condition, and two t-test on TA and NA items between conditions.

8.2.5.2.3. Analysis by Target-Alignment

Comparing performance on the surprise recognition test for TA and NA items between each group revealed that there was no significant difference in recognition rates for TA items between the ignored auditory words condition [$M = 0.84$, $SE = 0.06$] and ignored sounds condition [$M = 0.83$, $SE = 0.06$, $t(78) = 0.09$, $p = 0.928$], corroborated by a Mann-Whitney U Test [U = 0.09, $p = 0.928$]. However, NA items for ignored auditory words [$M = 0.82$, $SE = 0.03$] were recognized significantly less often than NA items for ignored sounds [$M = 0.91$, $SE = 0.02$, $t(78) = 2.42$, $p < 0.02$], which failed to meet Bonferroni corrected levels of significance ($p < 0.005$), corroborated by a Mann-Whitney U Test [U = 3.20, $p = 0.001$], which met corrected significance levels (see Figure 30).

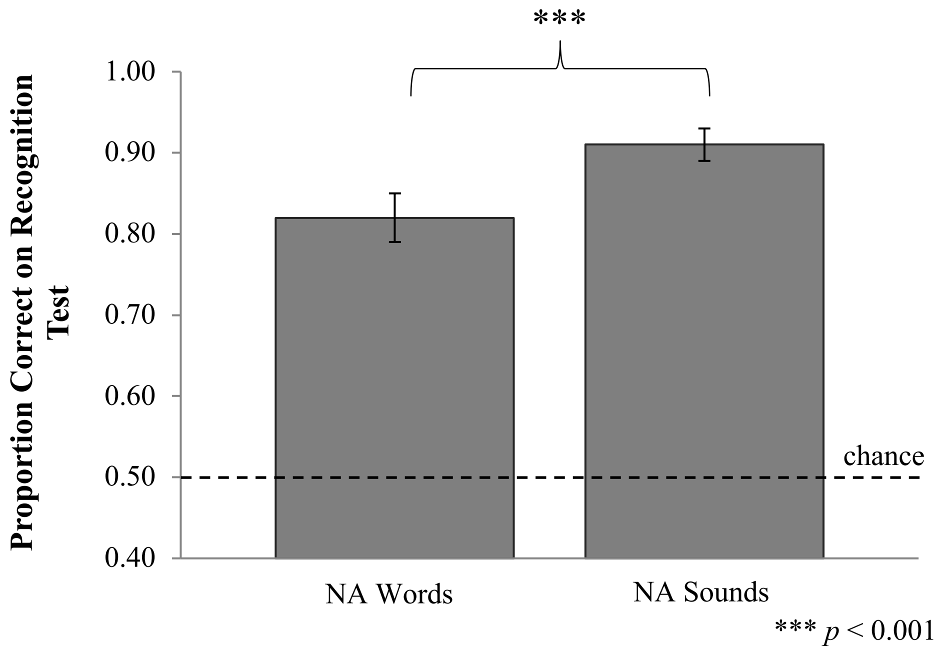


Figure 30. Recognition rates for NA items during the surprise recognition test in Experiment 2b. The primary task was presented in the visual modality, while ignored items were presented in the auditory modality. "NA Words" indicates recognition rates for non-aligned auditory words from the "Ignored Words" condition and "NA Sounds" indicates recognition rates for non-aligned sounds from the "Ignored Sounds" condition. Error bars represent the standard error for each variable. Both NA items were recognized at rates significantly above chance (0.50 – significance not shown). Those in the ignored sounds condition recognized significantly more NA items (M = 0.91, SE = 0.02) than those in the ignored words condition (M = 0.82, SE = 0.03, $p < 0.001$), suggesting that NA sounds were facilitated to a greater extent compared to NA words.

8.2.6. Discussion: Experiment 2b

Experiment 2b evaluated potential differences in the facilitated processing of previously unattended auditory lexical (i.e., words) and non-lexical (i.e., sounds) stimuli presented under cross-modal auditory/visual conditions in an attention-demanding task. In order to assess processing rates for attended and ignored items, between groups comparisons were made on primary task and surprise recognition test performance.

Regarding primary task performance, recall that participants were asked to attend only to the visual stream (i.e., pictures or written words) while ignoring the simultaneously presented auditory stream. Both groups obtained a proportion of hits significantly above chance, indicating that participants in both conditions were able to successfully identify visual targets in the cross-modal auditory/visual stream. There was no significant difference in d' or RT between groups. Thus, it appears that when the primary task is presented under cross-modal conditions, and participants monitored the visual stream for immediate repetitions, they were equally able to accurately identify and respond to targets, regardless of stimulus type (i.e., pictures or written words).

Analysis of the surprise recognition test for previously unattended auditory items (i.e., ignored auditory words or ignored sounds) revealed that both groups obtained an accuracy score (Hits + CRs) significantly above chance, again suggesting that both groups were able to successfully complete the task. There was no significant difference in d' between groups, which suggests that participants were comparable in their ability to distinguish between previously heard items (auditory words or sounds) and foil items, overall, during the surprise recognition test.

Considering the critical analysis on the surprise recognition test data, a two-factor (2x2) repeated measures ANOVA was conducted on participant responses with target-alignment (TA vs. NA) as the within subjects factor, and stimulus type (ignored sounds vs. ignored auditory words) as the between subjects factor. There was no main effect for target-alignment, suggesting that, overall, TA items (auditory words and sounds) were not recognized significantly more often than NA items (auditory words and sounds) during the surprise recognition test. There was also no main effect for stimulus type, indicating that previously ignored sounds (TA and NA) were not recognized significantly more often than previously ignored auditory words (TA and NA), and there was no interaction, suggesting that recognition rates for TA and NA items were not significantly different between previously ignored sounds and auditory words. Thus, it appears that, in the auditory modality, temporal and spatial alignment between an ignored auditory item and an attended visual target did not lead to facilitated processing of the ignored information above equally presented, unattended, yet temporally non-aligned items, and later recognition rates were not different between stimulus types.

It is possible that even under cross-modal conditions, the visual task was demanding enough to impart sufficient cognitive load, which may continue to mitigate the effects of target-alignment in the auditory modality. Indeed, reductions in neural activation in the auditory cortex associated with segregation of complex auditory stimuli have been observed when the sound stream is presented concurrently with an attended, high load, visual task (Molloy, Lavie, & Chait, 2019; Tillmann, & Swettenham, 2017), suggesting that visual load can limit the computational capacity of the auditory system. However, it should be noted that overall recognition rates for all item types (TA/NA and auditory words/sounds) were significantly above chance, which suggests that processing for the unattended auditory items was unilaterally facilitated during the primary task. The overall high recognition rates observed during the surprise recognition test may then be attributed to two main factors.

First, as mentioned before, a growing body of research has demonstrated that cross-modal presentation leads to improved task performance and facilitated processing for unattended items due to the availability of additional attentional resources (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984). Thus, while the cognitive load imparted by the visual task may have imposed some restrictions on the extent to which unattended auditory items were processed – perhaps limiting the facilitatory effect of target-alignment – the cross-modal presentation allowed for non-selective facilitation of all presented, yet unattended, auditory items.

Second, all unattended items were presented at a very high rate and an equal number of times (i.e., 120 presentations each, during the primary task). Therefore, the frequency of stimulus presentation, coupled with the additional attentional resources available when the primary task was presented cross-modally, may have allowed for all unattended items to be processed somewhat deeply within the cognitive system, leading to some form of representation being maintained in memory, resulting in overall high recognition rates during the surprise recognition test. The lack of difference between stimulus types is likely attributed to the activation of competitive representational nodes, both for auditory words and sounds (which are likely to contain semantic information and the precise semantic associations carried by each sound may be somewhat ambiguous, as is the case with words). Under this framework, competitive representational activation would cause interference either during encoding or retrieval for both types of information, which suggests similar processing procedures for both stimulus types, leading to similar recognition rates later, when the unattended information is presented to the auditory modality.

Despite the fact that an interaction was not observed, pre-planned analyses investigated the relationship between recognition rates for TA and NA items within and between each stimulus condition. For both conditions (ignored auditory words and ignored sounds), recognition rates for TA items and NA items were significantly above chance. There was also no significant difference in recognition rates between TA items and NA items within each stimulus type. Finally, comparing performance on recognition rates for TA and NA items *between* stimulus types revealed no significant difference in recognition for TA items between previously ignored sounds and auditory words, however, NA sounds were recognized significantly more often than NA auditory words by conventional standards, but not according to Bonferroni corrected standards. Taken together, it seems that target-alignment did, not play a critical role in facilitated processing for unattended information (regardless of stimulus type) in the auditory modality, even when initially presented under cross-modal conditions. As with Experiment 1b, non-lexical auditory stimuli may be processed in a manner quite similar to that of lexical auditory information when actively ignored, even under cross-modal conditions, leading to similar recognition rates during the surprise recognition test.

9. Discussion, Limitations, and Future Directions

9.1. Discussion

The overarching aim of this dissertation was to explore the extent to which processing for unattended lexical and non-lexical stimuli may be facilitated under unimodal and cross-modal conditions, when participants are otherwise engaged in an attention-demanding task. The body of work detailed in this doctoral dissertation represents a novel contribution to current scientific understanding regarding how the human cognitive system facilitates processing for different types of unattended information within a single sensory modality (vision *or* audition) and under cross-modal conditions (vision *and* audition). While much work has been done to investigate the role of stimulus presentation frequency and temporal alignment in the facilitated processing for unattended information (Dewald et al., 2013; Sasaki, Náñez, & Watanabe, 2010; Seitz & Watanabe, 2003, 2005; 2009; Walker et al., 2014; Walker et al., 2017; Watanabe et al., 2001), and consideration has been given to the saliency of the unattended stimuli (Tsushima et al., 2006, 2008), less attention has been given to the role of stimulus type itself or the sensory modality of stimulus presentation. While other veins of research have examined processing rates for lexical and non-lexical information in the visual modality under conditions of directed attention (Amit et al., 2009; Carr, McCauley, Sperber, & Parmelee, 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980) as well as when these items are actively ignored (Tipper, 1985; Tipper & Driver, 1988) fewer have considered the auditory modality under attended (Vitevitch & Luce, 1998, 2016) and ignored conditions (Koreimann et al., 2014), and none – at the time of writing this doctoral dissertation – have systematically explored the extent to which unattended auditory or visual, lexical and non-lexical information, may be facilitated under cross-modal conditions.

To better direct the discussion of the findings presented here, it is important to reiterate the two main aims that were first presented in chapter six:

- **Aim 1** was to evaluate the role of stimulus type (i.e., lexical vs. non-lexical information) in the facilitated processing of task-irrelevant stimuli
- **Aim 2** was to evaluate the role of multisensory presentation in the facilitated processing of task-irrelevant stimuli

Each aim was addressed through a series of two experiments. Below, table 1 summarizes each experiment, conditions, and observed results from the critical analyses of the surprise recognition test. Specifically, Experiment 1 compared facilitation rates for unattended lexical and non-lexical stimuli (i.e., written/auditory words and pictures/sounds, respectively) under unimodal visual (E1a) and unimodal auditory conditions (E1b). Taking the investigation a step further, Experiment 2 compared facilitation rates for unattended lexical and non-lexical stimuli under cross-modal conditions wherein participants either ignored lexical or non-lexical *visual* information (written words and pictures, respectively) while attending to an auditory stream (E2a), or ignored lexical or non-lexical *auditory* information (auditory words and sounds, respectively) while attending to a visual stream (E2b). In order to streamline the discussion of the experiments presented thus far, a brief summary of the primary findings will be reviewed for each experiment, followed by an exploration of the broad pattern of results observed across all four manipulations.

Table 1. A Summary of Results from the Critical Analyses		
Aims / Experiment	**Conditions**	**Results**
Aim 1 Experiment 1a (Unimodal Visual Conditions)	Attend Written Words / Ignored Pictures Attend Pictures / Ignored Written Words	**Main Effects** TA > NA Pictures > Written Words **Interaction** N/A
Aim 1 Experiment 1b (Unimodal Auditory Conditions)	Attend Auditory Words / Ignored Sounds Attend Sounds / Ignored Auditory Words	**Main Effects** N/A **Interaction** N/A
Aim 2 Experiment 2a (Cross-Modal Visual Conditions)	Attend Auditory Words / Ignored Pictures Attend Sounds / Ignored Written Words	**Main Effects** TA > NA Pictures > Written Words **Interaction** N/A
Aim 2 Experiment 2b (Cross-Modal Auditory Conditions)	Attend Written Words / Ignored Sounds Attend Pictures / Ignored Auditory Words	**Main Effects** N/A **Interaction** N/A

9.1.1. Experiment 1a: Purpose and Key Findings

Experiment 1a evaluated the role of stimulus type in the facilitation of unattended lexical (i.e., words) and non-lexical (i.e., pictures) items under unimodal visual conditions in an attention-demanding task.

During the primary task, participants were shown a rapidly presented compound visual stream containing pictures and words and were required to attend to one dimension (e.g., pictures) for immediate repetitions, while ignoring the other dimension (e.g., words). While both groups had a proportion of hits significantly above chance, those in the attend pictures condition a significantly lower d’ compared to those in the attend words condition. These finding suggest that, while both groups were successful at completing the primary task, when participants were attending to pictures they were more likely to respond even when a target was not present.

Regarding the critical analyses for the surprise recognition test, the ANOVA revealed a main effect for target-alignment, suggesting, overall, that TA items (words and pictures) were recognized more often compared to NA items. This finding replicates previous work by Dewald et al. (2013, see also Walker et al., 2014, 2017) under unimodal visual conditions, but with a much larger sample size, and demonstrates the robustness for the role of temporal-alignment in the facilitated processing of unattended visual information. There was also a main effect for stimulus type, indicating that previously ignored pictures were recognized significantly more often than previously ignored words, with post-hoc analyses revealing that this was largely driven by recognition performance for NA pictures, which were recognized significantly more often compared to NA words. Thus, it appears that the observed differences in processing for attended pictures and words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980) extends to conditions in which these items are actively ignored in an attention-demanding visual task. Finally, there was no interaction, suggesting that recognition rates for TA and NA items were not different between stimulus types.

Together, the results suggest that when presented under unimodal visual conditions, unattended pictures are processed more extensively than unattended words, leading to higher recognition rates during the surprise recognition test. Furthermore, target-alignment boosts processing for TA items, preferentially, and this process appears to proceed in a similar manner for both pictures and words despite these items being processed somewhat differently, in general, by the cognitive system.

9.1.2. Experiment 1b: Purpose and Key Findings

Experiment 1b evaluated the role of stimulus type in the facilitation of unattended lexical (i.e., words) and non-lexical (i.e., sounds) items under unimodal auditory conditions in an attention-demanding task.

During the primary task, participants listened to a rapidly presented compound auditory stream containing sounds and auditory words and were required to attend to only one dimension (e.g., sounds) for immediate repetitions, while ignoring the other (e.g., auditory words). While both groups had a proportion of hits significantly above chance, those in the attend sounds condition a significantly lower d' compared to those in the attend auditory words condition. This finding suggests that, while both groups were successful at completing the primary task, participants attending to sounds were more likely to respond even when a target was not present.

The critical analyses of the surprise recognition test revealed no main effect for target-alignment, suggesting that TA items (auditory words and sounds) were not recognized significantly more often than NA items. There was also no main effect for stimulus type, indicating that previously ignored sounds were not recognized significantly more often that previously ignored auditory words, and there was no interaction between these two dimensions, suggesting that recognition rates for TA and NA items were not different between the two stimulus types.

Together, the results suggest that when presented under unimodal auditory conditions, unattended sounds are not processed more extensively than unattended auditory words. Finally, target-alignment does not appear to play a critical role in facilitated processing for unattended auditory information (regardless of stimulus type).

9.1.3. Experiment 2a: Purpose and Key Findings

Experiment 2a evaluated the role of stimulus type in the facilitation of unattended visual lexical (i.e., written words) and non-lexical (i.e., pictures) items under cross-modal auditory/visual conditions wherein participants perform an attention-demanding auditory task.

During the primary task, participants were presented with a rapid compound auditory/visual stream containing auditory lexical (spoken words) or non-lexical (sounds) items simultaneously with visual non-lexical (pictures) or lexical (written words) items, respectively. Participants were required to attend only to the auditory dimension (e.g., sounds) for immediate repetitions, while ignoring the visual dimension (e.g., written words). While both groups had a proportion of hits significantly above chance, those in the attend sounds condition had a significantly *higher* d', and faster RTs to identified targets, compared to those in the attend auditory words condition. These findings suggest that when participants were monitoring the auditory stream for immediate repetitions while ignoring a simultaneously presented visual stream, sound targets were identified, and responded to, more readily compared to auditory word targets.

Overall recognition rates for previously ignored visual words and pictures were well above chance. The critical analyses of the surprise recognition test revealed a main effect for target-alignment, suggesting that TA items (words and pictures) were recognized significantly more often than NA items. This finding demonstrates the robustness of the Dewald et al. (2013) study, and extends the role of temporal-alignment in the facilitated processing of unattended visual information to cross-modal conditions. There was also a main effect for stimulus type, indicating that previously ignored pictures were recognized significantly more often than previously ignored words. Thus, it appears that the observed differences in processing for attended pictures and words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980) also extends to conditions in which the unattended visual information is initially presented with an attended auditory stream in an attention-demanding cross-modal task. Finally, there was no interaction, suggesting that recognition rates for TA and NA items were not different between stimulus types. As with Experiment 1a, this finding suggests that facilitation by target-alignment proceeds in a similar manner between stimulus types, despite pictures being processed more readily than words.

Collectively, the results suggest that when presented under cross-modal auditory/visual conditions, all unattended visual items are processed extensively, leading to high recognition rates for both stimulus types. However, unattended pictures continue to be processed more extensively than unattended words (as seen under unimodal visual conditions in Experiment 1a), leading to higher recognition rates during the surprise recognition test. Furthermore, target-alignment continues to boosts processing for TA items, preferentially, under cross-modal conditions (again, as seen under unimodal visual conditions in Experiment 1a), and this process appears to proceed in a similar manner for both pictures and words despite these items being processed somewhat differently, in general, by the cognitive system.

9.1.4. Experiment 2b: Purpose and Key Findings

Experiment 2b evaluated the role of stimulus type in the facilitation of unattended auditory lexical (i.e., auditory words) and non-lexical (i.e., sounds) items under cross-modal auditory/visual conditions wherein participants perform an attention-demanding visual task.

During the primary task, participants were presented with a rapid compound auditory/visual stream containing visual lexical (written words) or non-lexical (pictures) items simultaneously with auditory non-lexical (sounds) or lexical (auditory words) items, respectively. Participants were required to attend only to the visual dimension (e.g., pictures) for immediate repetitions, while ignoring the auditory dimension (e.g., auditory words). Both groups obtained a hit rate significantly above chance and there was no significant difference in d' between groups. Thus, when participants were monitoring the visual stream for immediate repetitions while ignoring a simultaneously presented auditory stream, they were able to successfully complete the task and they were equally able to accurately identify and respond to task-relevant targets, regardless of stimulus type (i.e., pictures or written words).

While overall recognition rates for previously ignored auditory words and sounds were well above chance, the critical analyses of the surprise recognition tests for Experiment 2b revealed no main effect for target-alignment, suggesting that TA items (auditory words and sounds) were not recognized significantly more often than NA items. There was also no main effect for stimulus type, indicating that previously ignored sounds were not recognized significantly more often than previously ignored auditory words, and there was no interaction between these two dimensions, suggesting that recognition rates for TA and NA items were not different between the two stimulus types.

Together, the results suggest that when presented under cross-modal auditory/visual conditions, all unattended auditory items (i.e., sound and auditory words) are processed somewhat extensively, leading to overall high recognition rates. However, there does not appear to be a significant difference in facilitated processing for unattended sounds compared to auditory words. Finally, target-alignment does not appear to play a critical role in facilitated processing for unattended auditory information, even under cross-modal conditions, regardless of stimulus type (as seen in Experiment 2b under unimodal auditory conditions).

9.1.5. General Discussion

While it should be noted that direct comparisons between experiments are not possible, the critical analyses across all four experiments revealed an interesting pattern of results. Figure 31 depicts findings from the critical analyses of each experiment together. Because the results of each experiment were quite different depending on the sensory modality of the ignored items (lexical and non-lexical), relevant results from the critical analyses of Experiment 1a and Experiment 2a (i.e., ignored visual information) will be discussed together, followed by Experiment 1b and Experiment 2b (i.e., ignored auditory information).

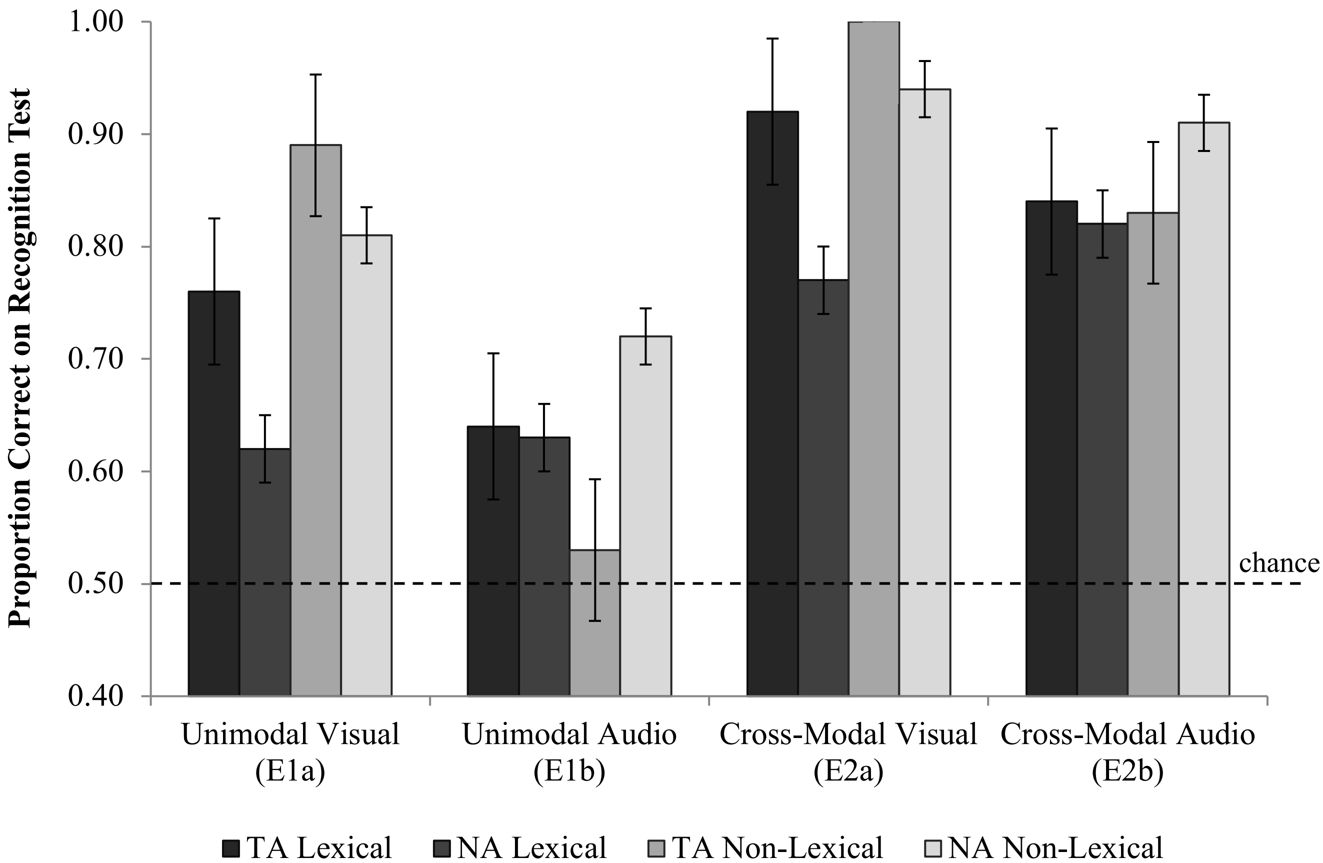


Figure 31. Results from the critical analyses on surprise recognition test performance for previously ignored information across all four experiments divided by target-alignment (TA vs. NA) and lexical status (lexical vs. non-lexical) for each presentation condition. The two darkest bars represent the lexical items (TA and NA, respectively), while the two lightest bars represent the non-lexical items (TA and NA, respectively). E1a: participants were presented all items (attended and ignored) under unimodal visual conditions. E1b: participants were presented all items (attended and ignored) under unimodal auditory conditions. E2a: participants attended the auditory stream while ignoring the visual stream. E2b: participants attended the visual stream while ignoring the auditory stream.

Experiment 1a (E1a) represents a replication of previously conducted research by Dewald et al. (2013) under unimodal visual conditions, but with a much larger sample size; Experiment 2a (E2a) extends the investigation of facilitated processing for unattended visual information to include cross-modal auditory/visual presentation of the attended (auditory) an unattended (visual) stimuli during the primary task. For both experiments (E1a and E2a), the critical analyses revealed a main effect for target-alignment, suggesting that TA visual items (words and pictures) were recognized significantly more often than NA items. This finding demonstrates the robustness of the Dewald et al. (2013) study, and provides further support for the role of temporal-alignment in the facilitated processing of unattended visual information regardless of unimodal (E1) or cross-modal (E2a) presentations and stimulus type (i.e., lexical or non-lexical information). There was also a main effect for stimulus type, indicating that previously ignored pictures were recognized significantly more often than previously ignored words. Thus, it appears that the observed differences in processing for attended pictures and words (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980) extends to conditions in which these items are actively ignored in an attention-demanding visual task (E1a) and when the unattended visual information is initially presented with an attended auditory stream in an attention-demanding cross-modal task (E2a). Finally, there was no interaction, suggesting that recognition rates for TA and NA items were not different between stimulus types.

Taken together, findings from E1a and E2a lend further support to the idea that pictures are processed more readily than words, which may be attributed to the variant properties of each stimulus type. Under the current framework (Atkinson & Shiffrin, 1968, 1971; Baddeley, 1992, 2012; Baddeley & Hitch, 1974; Hitch & Baddeley, 2017; Huitt, 2003; Shiffrin & Atkinson, 1969; Skóra & Wierzchoń, 2016; Sørensen, 2017; Wolfe, 2001; Wolfe & Gray, 2007), the observed facilitation for pictures compared to words in the visual modality may be reliant on holistic processing of the visual features for each item and their associated semantic links in LTM. Due to the lower level of conceptual abstraction resulting from the concrete representation of the item itself, pictures may bypass some of the processing constraints imposed on written words. This is likely because a word must first be decoded (i.e., they must be read) before the semantic associations can be retrieved. While a selection of investigations assert that reading is a largely automatic process (Augustinova & Ferrand, 2014, Deutsch & Deutsch, 1963; LaBerge & Samuels, 1974; MacLeod, 1991, Posner, 1975), others suggest that this may not always be the case and that reading comprehension may be tied, at least in part, to attentional processes (Besner, Risko, & Sklair, 2005; Rees et al., 1999; Reynolds & Besner, 2006; Waechter, S., Besner, & Stolz, 2011).

Despite divergent perspectives on the extent to which reading occurs automatically, it is broadly agreed upon that word processing may progress through a series of stages *before* semantic evaluation can occur. For example, LaBerge and Samuels (1974) suggest that reading begins with assessment of the visual features, followed by an evaluation of the phonological properties associated with the stimulus, then activation of relevant episodic memories associated with the word itself, and finally contextual and semantic evaluation. Indeed, some studies have proposed that because pictures are concrete representations of items, they do not require the same level of decoding that words do (Amit et al., 2009; Carr et al., 1982; Hogaboam & Pellegrino, 1978; Smith & Magee, 1980). While it is probable that processing for pictures, like words, begins with an evaluation of relevant visual features such as size, shape, and color, pictures may not be evaluated for phonological properties to the same extent as words. This stage may be bypassed, at least in part, resulting in earlier and more extensive memory activation and semantic evaluation. Indeed, a sample of studies suggest that pictorial stimuli result in greater amounts of cognitive elaboration, compared to words, leading to a greater number of storage locations within LTM as well as a more diverse network of related semantic connections, which increases the likelihood of that information being retrieved at a later time (Adaval, Saluja, & Jiang, 2019; Edell & Staelin, 1983). While it should be noted that words also contain diverse connections within the semantic network, a greater portion of them may be competitive (rather than related), due to words having a higher level of conceptual abstraction, leading to interference (Vitevitch & Luce, 1998, 2016). Thus, even when processed extensively (i.e., evaluated for semantic content and stored in LTM) words may still have additional levels of constraint imposed on them compared to pictures.

The main effect for target-alignment suggests that the facilitatory effect of temporal pairing between an unattended visual item and an attended task target (either visual or auditory) is indeed robust, leading to higher recognition rates for unattended TA visual items compared to unattended NA visual items. Interestingly, though main effects for stimulus type and target-alignment were observed in both E1a and E2a, neither experiment revealed a significant interaction between these two dimensions. The lack of an interaction indicates that facilitation by target-alignment does not proceed differently for lexical and non-lexical visual items, despite the observed differences in processing for each type of stimulus. Thus, it is quite probable that facilitation by target-alignment operates, primarily, on a *temporal* dimension (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe, 2001). Presenting an unattended item with an attended task-relevant target in the visual modality may result in some degree of signal integration between the attended and unattended stimuli. Essentially, the unattended information may "come along for the ride" as the task-relevant target is identified and response operations are generated, akin to the notion of lag 1 sparing observed in the attentional blink literature (Chun & Potter, 1995; Shapiro, 1994; Wolfe & Gray, 2007). This would suggest that target-alignment may have overlapping principles as those observed in attentional blink paradigms (Chun & Potter, 1995; Shapiro, 1994; Wolfe & Gray, 2007) suggesting that the unattended item is not filtered out at the response selection bottleneck in the system (Wolfe & Gray, 2007), but passed through along with the target and is subsequently stored, at least to some extent, in LTM, thereby facilitating recognition of the previously ignored items when encountered later, compared to NA items.

While the complementary findings between E1a and E2a were extensive, a notable difference between these two studies is also worth mentioning. Despite the fact that a direct statistical comparison between experiments cannot be made, the numerical trend suggests that recognition rates for both previously ignored stimulus types (words and pictures) may be higher when initially presented under cross-modal (E2a) compared to unimodal conditions (E1a). Recognition rates for all previously ignored visual items (lexical and non-lexical) were above chance in both experiments, however participants in E1a recognized 64% of previously ignored words and 82% of previously ignored pictures, while those in E2a, recognized 79% of previously ignored words and 95% of previously ignored pictures. This trend aligns well with the idea that presenting the primary task cross-modally reduces cognitive load and frees up attentional resources between the auditory and visual sensory systems (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984) allowing for greater amounts of unattended visual information to be processed regardless of stimulus type. However, unattended pictures still appear to be processed more extensively compared to unattended words under cross-modal conditions, leading to higher recognition rates during the surprise recognition test.

Overall, the results from E1a successfully replicate previous work demonstrating that presenting an unattended visual item with an attended visual target leads to facilitated processing for the target-aligned information compared to equally presented, but non-aligned, information (see Dewald et al., 2013; Seitz & Watanabe, 2003, 2005; Walker et al., 2014, 2017), and E2a extends these findings to conditions in which the unattended visual information is presented with an attended auditory target. These studies also appear to confirm the hypothesis that pictures are processed more readily than words both when attended and when actively ignored as well as when presented under unimodal visual conditions and under cross-modal conditions. Finally, results suggest that the mechanism responsible for facilitated processing of unattended target-aligned, compared to non-aligned, visual information operates uniformly across lexical and non-lexical stimuli. Next, consideration will be given to conditions in which the unattended information is presented to the auditory modality under unimodal conditions (i.e., Experiment 1b) and under cross-modal conditions (i.e., Experiment 2b).

Experiment 1b (E1b) extended investigations from Dewald et al. (2013) by examining the extent to which processing for lexical and non-lexical stimuli may be facilitated under unimodal auditory conditions. Experiment 2b (E2b) extended these investigations further to include cross-modal auditory/visual presentation of the attended (visual) and unattended (auditory) stimuli during the primary task. The critical analyses of the surprise recognition tests for both experiments (E1b and E2b) revealed no main effect for target-alignment, suggesting that TA items (auditory words and sounds) were not recognized significantly more often than NA items. There was also no main effect for stimulus type, indicating that previously ignored sounds were not recognized significantly more often than previously ignored auditory words, and there was no interaction between these two dimensions, suggesting that recognition rates for TA and NA items were not different between the two stimulus types.

There are several factors that may contribute to the divergent findings between unattended visual (E1a and E2a) and auditory conditions (E1b and E2b). First, the lack of a main-effect for target-alignment may be attributed, in part, to the unique characteristics of presenting information to the auditory modality. Recall that processing for auditory information relies heavily on temporal aspects of the stimuli, such as onset, duration, amplitude, frequency, and intonation (Henny et al., 2018; Notter et al., 2018; Simon & Winkler, 2018; Snyder, 2015; Teki et al., 2016; Teng et al., 2017; Tóth et al., 2016; Wang et al., 2018), while the visual modality is more reliant on the featural composition of the stimuli, such as color, line, shape, size, and spatial location (Bundesen & Pedersen, 1983; Gilbert & Li, 2013; Livingstone, & Hubel, 1988; Morey, 2018; Poort et al., 2012; Self et al., 2013; Shipp & Zeki, 1985; Treisman, 1982; Wagemans et al., 2012; Wolfe, 1994). In E1b the temporal aspects of the concurrent sound streams were held constant between the attended and ignored dimensions. That is, when the primary task was presented under unimodal auditory conditions, both auditory streams (lexical and non-lexical) were played concurrently and, importantly, matched on temporal dimensions such as onset, duration, and amplitude, which has been demonstrated to increase the likelihood for signal integration in the auditory modality (Simon & Winkler, 2018). Thus, the controlled experimental conditions used in E1b may result in a higher level of signal integration between the attended and unattended streams compared to unimodal visual conditions (E1a). This is an important consideration, given that facilitation by target-alignment appears to operate primarily on a temporal dimension (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe, 2001). Arguably, presenting all items in the auditory modality would cause both pieces of information (attended and ignored) to advance through the cognitive system in parallel regardless of target-alignment, resulting in global facilitation for all unattended items rather than preferential facilitation for TA items over NA items.

Second, a higher level of signal integration between the attended and ignored auditory streams would make signal segregation more difficult, meaning participants may have had a harder time cognitively separating the attended information from the unattended information in E1b. This would increase cognitive load associated with the primary task (Cartwright-Finch & Lavie, 2007; Lavie, 2005, 2010) and lead to interference within the system, which may limit the extent to which each individual item is processed and subsequently stored in LTM. Again, direct statistical comparisons between experiments are not possible, however an evaluation of the numerical trends between E1a and E1b do lend support for this hypothesis. Overall, participants in E1a (i.e., unimodal visual conditions) recognized 83% of TA items and 72% of NA items, while participants in E1b (i.e., unimodal auditory conditions) recognized 58% of TA items and 68% of NA items. Thus, in addition to no significant difference in recognition rates between TA and NA items, overall recognition rates for participants in E1b suggest lower performance on the surprise recognition test compared to participants in E1a.

However, this does not explain why a main effect for target-alignment was not observed in E2b, when the primary task was presented under cross-modal conditions. Because E2b presented the attended stream to the visual modality and the unattended stream to the auditory modality, the potential for signal integration between these two dimensions was much lower compared to the unimodal conditions of E1b. While many studies suggest that cross-modal presentation leads to reductions in cognitive-load and improved task performance (Driver & Spence, 1998, 2004; Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco et al., 2005; Wickens, 1984), recent investigations also suggest that high-load visual tasks can impair computational capacity and limit signal segregation for a concurrently presented auditory stream (Molloy et al., 2019; Tillmann & Swettenham, 2017).

Interestingly, numerical trends do suggest that participants in E2b demonstrated an improvement on surprise recognition test performance compared to E1b. Overall, participants in E2b (i.e., cross-modal conditions) recognized 84% of TA items and 86% of NA items while participants in E1b (i.e., unimodal conditions) recognized 58% of TA items and 68% of NA items. Therefore, it appears that cross-modal presentation of the attended and unattended stimuli in E2b significantly reduced cognitive-load associated with the primary task, thereby freeing up attentional resources, leading to global facilitation for the unattended auditory items. However, the primary task still demands a high level of sustained visual attention in order to successfully identify and respond to targets in the RSVP stream. Dovetailing with this notion, it is worth noting that a series of studies suggest that humans are visually dominant (Colavita, 1974; Colavita, & Weisberg, 1979; Posner et al., 1976; Sinnett, Spence, & Soto-Faraco, 2007; Spence, 2009; Spence, Parise, & Chen, 2012), indicating that we tend to rely on our visual sense and, as a result, divert more of our attentional resources toward processing visual information. This may place some level of constraint on the extent to which the unattended auditory items are processed (Molloy et al., 2019; Tillmann & Swettenham, 2017) and limit the facilitatory effect of target-alignment in the auditory modality, even under cross-modal conditions.

The lack of a main effect for stimulus type in E1b and E2b was surprising. Research from dichotic listening tasks demonstrates that even ignored lexical information may still be evaluated for semantic content, at least in part, (Cherry, 1953; Conway, Cowan, & Bunting, 2001; Dalton & Fraenkel, 2012; Giraudet, St-Louis, Scannella & Causse, 2015; Haykin & Chen, 2005; Macdonald & Lavie, 2011; Murphy & Greene, 2015; Raveh & Lavie, 2015; Treisman, 1960). Taken together, it was reasonable to hypothesize that a main effect for stimulus type would be observed under the current conditions. All presented words were intentionally selected for their high frequency in the English language. Thus, when encountered, these words would activate a large number of representational competitors that may have incompatible semantic links, increasing interference in the cognitive system, and limiting processing for those items (Vitevitch & Luce, 1998, 2016). If unattended lexical information is still evaluated for semantic content, as suggested by the dichotic-listening literature, these constraints should apply even when these items are actively ignored, leading to lower recognition rates compared to unattended sounds.

However, there is a notable difference between the sound stimuli from the current experiments and those used by Vitevitch and Luce (1998), which may provide an explanation for the lack of a main effect for stimulus type. Recall that the sounds used in E1b and E2b were selected to match the relative frequency of the chosen words and pictures. That is, the non-lexical stimuli employed here represented a variety of common, naturally occurring, sounds in an everyday environment, such as horns, bells, animal noises, sneezes, and belches, while Vitevitch and Luce (1998) presented participants with sub-lexical speech sounds such as /k/, /t/, or /æ/. The extent to which environmental sounds are processed differently from auditory linguistic sounds is not well investigated. Furthermore, there is little literature available on whether or not environmental sounds contain semantic content or activate semantic representations within our cognitive system to the same extent that linguistic sounds do. However, Ballas and Howard (1987) posit that environmental sounds may become integrated into the cognitive system based on properties that are similar to those used in speech perception, which would suggest they may be treated in a similar fashion, and a more recent investigation appears to mirror this sentiment when these items are ignored (Koreimann et al., 2014). On the other hand, neural segregation has also been observed for processing of speech sounds and non-speech sounds (Binder, Frost, Hammeke, Bellgowan, Springer, Kaufman, & Possing, 2000), suggesting at least some level of separation.

Thus, the sounds utilized in the current experiments, arguably, contain some level of semantic content that may be somewhat ambiguous. Meaning that environmental sounds may maintain a level of conceptual abstraction that is more comparable to words than to pictures. For example, hearing a sound (like a horn) may activate divergent links to potentially competitive representational nodes – indeed the sound of a horn may be tied to a large number of contextual and semantic associations that are not necessarily compatible – thereby causing interference. As a result environmental sounds may be processed in a manner analogous to that of spoken words and this may proceed even when unattended. This is compounded by the fact that both stimulus types are presented at a very high rate during the primary task. Recall that the primary task consists of 960 trials for which there are only seven unattended NA items and one unattended TA item, all presented with equal frequency (i.e., 120 times each). Thus, despite attention being directed elsewhere, the high rate of presentation may lead to these items being processed extensively as participants complete the attended primary task. This would increase the likelihood that all unattended auditory items (i.e., words and sounds) are evaluated at the semantic level, leading to the activation of competitive representational nodes in the semantic network. This could potentially result in comparable levels of interference for both stimulus types, resulting in similar recognition rates during the surprise recognition test.

Again, E1b and E2b revealed complementary findings for processing unattended auditory information between unimodal and cross-modal conditions. However, there is a notable difference worth discussing. While both experiments revealed no main effects and no interaction, numerical trends from the surprise recognition tests suggest improved performance in E2b compared to E1b. Overall, participants in E1b (i.e., unimodal auditory conditions) recognized 64% of previously ignored auditory words and 70% of previously ignored sounds, while participants in E2b (i.e., cross-modal conditions) recognized 83% of previously ignored auditory words and 87% of previously ignored sounds. The cross-modal presentation of the primary task in E2b may account for the overall higher recognition rates for previously unattended items. As noted earlier, presenting all items in the unimodal auditory stream, in E1b, likely leads to higher instances of signal integration, thereby lowering overall recognition rates later. However, the cross-modal presentation in E2b removes this confound allowing for each stimulus stream to be processed somewhat separately, arguably freeing attentional resources and allowing for greater amounts of unattended auditory information to be processed, leading to overall higher recognition rates under cross-modal conditions while still maintaining similar recognition rates between stimulus types.

Overall, the results from E1b demonstrate that target-alignment does not play a critical role in facilitated processing for unattended auditory information (regardless of stimulus type) under unimodal conditions and E2b confirms this under cross-modal conditions. This may be attributed to greater amounts of signal integration between the attended and ignored auditory stream under unimodal conditions in E1b, which is likely to increase cognitive load associated with the primary task and place additional constraints on the extent to which ignored auditory items are processed, regardless of target-alignment. This may continue to be the case under conditions of cross-modal presentation, as the visual task in E2b is still quite demanding and participants may divert more of their attentional resources toward processing the visual stream despite an overall reduction in cognitive load. Furthermore, it is likely that the lexical and non-lexical auditory stimuli used here may have more semantic overlap than was initially anticipated. Despite being actively ignored in E1b and E2b, the auditory information may still be processed extensively due to the high rate of presentation during the primary task, allowing for similar recognition rates later.

Thus, it appears that processing for unattended lexical and non-lexical information may proceed differently depending on the sensory modality in which those items are presented and the precise type of stimulus being utilized. Pictures appear to be processed more extensively compared to written words, leading to a richer representation, as well a diverse set of complementary storage locations and retrieval pathways within LTM, thereby facilitating recognition for these items later. However, unattended sounds and auditory words may be processed in a similar manner, ultimately leading to statistically indistinguishable recognition performance. Thus, auditory semantic access may be more indirect for both lexical and non-lexical information thereby limiting the extent to which these items may be stored in LTM, and recognized later compared to visual conditions.

The results of these experiments have important implications for those interested in utilizing words, pictures, and sounds to study information processing across sensory domains in an experimental paradigm. Ultimately, more research is required to understand the precise mechanisms driving the extent to which unattended lexical and non-lexical information may be processed in the visual and auditory modalities. Future studies should strongly consider the potential semantic content of non-linguistic auditory information when making comparisons between lexical and purportedly non-lexical stimuli, as there appears to be a great amount of overlap in how semantic auditory information may be processed, regardless of lexical status. Finally, while target-alignment does not appear to play a critical role under the current conditions, considerations should be given to the temporal dimensions of the auditory stimuli (under unimodal and cross-modal conditions), the various propensities for signal integration between vision and audition under compound stimulus presentations, and the role of visual dominance in attention-demanding tasks.

9.2. Limitations

There are several notable limitations of the current project. Frist, recall that the critical analysis on the surprise recognition tests examine recognition rates between the previously ignored TA items against recognition rates for the seven previously ignored NA items via a two-factor (2x2) repeated measures ANOVA. Furthermore, recognition for each item is measured on a binary, ordinal scale, wherein the correct identification of a previously presented item is coded as a ‘1’ while a miss is coded as a ‘0’. Response outcomes for the seven NA items are averaged resulting in a quasi-continuous variable that is then compared to the binary response outcome for the TA items. This data configuration violates the assumption that the dependent variable is measured at the continuous level, which limits the reliability of the test. Despite this fact, it should be noted that the employment of an ANOVA on categorical outcome variables is relatively wide-spread in the behavioral sciences (Jaeger, 2008) despite the complications that arise when doing so.

The major limitation involved when including binary data in an ANOVA centers around the fact that ANOVAs make comparisons between raw means. The use of raw means in an ANOVA is limited with binary data as the model may include a range of probable outcomes beyond the confines of a binary response. For example a 95% confidence interval may predict a range of possible means that includes an upper boundary above one, which is difficult to interpret as response proportions within binary data are confined between zero and one. Thus, the inclusion of binary data may result in the model accounting for probability mass among outcomes that are not possible under dichotomous response options, which may lead to an underestimation of the probability mass for response outcomes that are actually possible (Jaeger, 2008). Ultimately this may lead to spurious null and significant results that go beyond what is expected among typical type I and type II error rates. A mix-effect logistic regression (MELR) may be a more appropriate form of analysis to conduct as it allows for the inclusion of binary data by calculating log-odds rather than comparing raw means (Agresti, 2002; Agresti & Kateri, 2011; Chatterjee, & Hadi, 2015; Chatterjee, Hadi, & Price, 2000; Harrell, 2001, 2015; Jaeger, 2008), but are often not included in analyses.

In order to address this concern, a MELR was attempted on the data collected for this dissertation. However, it is important to note that the outcome variables being compared in the current studies do not conform to strict binary constructs. That is, while the collected responses for TA items are defined on a strictly binary scale (i.e., zero and one), the NA items are aggregated into a quasi-continuous variable, which makes interpreting the results of a MELR quite difficult. Indeed, multiple attempts to perform the analysis on the dissertation data were made with mixed levels of success and often resulting in overly low estimations of means. While every effort was made to ascertain the cause of this outcome, and identify the most appropriate modeling approach given the observed results, these efforts were largely unfruitful. It is possible that the statistical software used to conduct the analyses (i.e., SPSS V25) is not currently flexible enough to execute this type of analysis with these data. Future efforts to resolve this, potentially using more flexible and powerful software (e.g., R) will be undertaken.

Next, the cross-modal conditions did not include additional manipulations that may be worth considering. Specifically, the experiments here do not include a manipulation wherein participants attend to visual lexical items while ignoring auditory lexical items and vice versa, nor does it include a manipulation wherein participants attend to visual non-lexical items while ignoring auditory non-lexical items and vice versa. These manipulations were intentionally left out of the experimental design. Recall that previous studies utilizing this type of paradigm and stimuli have employed words and pictures or words and sounds (Dewald et al., 2013; Rees et al., 1999; Sinnett et al., 2006; Walker et al., 2014, 2017). Because this work is largely exploratory, the goal was to keep the experimental design as close to previous investigations as possible. Furthermore, presenting participants with the same type of information (i.e., lexical or non-lexical) across different sensory modalities may introduce additional factors that need to be taken into consideration when implementing the experimental design. For example, pictures and sounds maintain highly divergent features (as was outlined in chapter five and throughout the general discussion). On the other hand, despite the fact that lexical information may be handled somewhat differently between the visual and auditory modality, the amount of featural overlap remains quite high. While this presents practical obstacles (i.e., presenting two visually overlapping words may make individual stimulus identification quite difficult) including a lexical / lexical manipulation may add additional confounds to the experimental design. Namely, processing of lexical information results in the activation of competing representational nodes, whereas this does not appear to be the case for pictures as the level of conceptual abstraction is relatively low for this stimulus type. The extent to which this holds true for non-lexical sounds is subject to debate. Thus, it is highly likely that presenting compound audio-visual lexical streams may lead to high rates of inhibitory processing for both the attended and ignored streams as competing representation nodes become activated. This would represent a significant confound for the exploration of facilitated processing among task-irrelevant streams of information. However, future exploration will certainly consider these additional conditions in great detail, as it is a natural extension to the current work.

Another limitation worth addressing revolves around the observed differences between auditory and visual stimuli used in the current study. As noted several times throughout this dissertation, presenting information to the auditory modality involves characteristics quite different from those of the visual modality. The temporal constraints placed on processing for auditory information may leave this modality somewhat disadvantaged compared to vision, especially when the information is presented under unimodal conditions. This raises questions as to the appropriateness of comparing the results of these experiments to each other. Recall that direct statistical comparisons were not made, however speculations were provided based on observed differences in numerical trends between experiments. Due to the exploratory nature of this work, outlining the differences in observed response patterns between the conducted experiments provides a useful framework by which future studies may be conducted, despite the fact that direct statistical comparisons cannot be made.

Finally, it is also worth noting that the amount of available literature on processing for environmental sounds compared to auditory words is quite limited. As a result, the auditory stimuli used for the current studies were selected in order to align as closely with previous studies using this same paradigm as possible (Dewald et al., 2013; Sinnett et al., 2006; Walker et al., 2014, 2017). Specifically, the sounds were selected because they were believed to have characteristic that were largely comparable to those of pictures. That is, the pictures used in the current paradigm represented common, everyday objects, such as animals, tools, clothing, food, etc., and the sounds represented common, every day noises, such as horns, beeps, animal sounds, and non-lexical sounds generated by humans. While not explicitly lexical in nature (i.e., they are not presented as words), it is possible that the sounds used here have stronger semantic-lexical ties than the pictures do and this may limit the extent to which comparisons between these stimulus types can be made. Future attempts should focus on defining the extent to which purportedly non-lexical stimuli in both modalities may have overlapping properties with their lexical counterparts, as this may impact how these items are processed.

9.3. Future Directions

Despite the noted limitations, this vein of research represents an important contribution to our understanding of human cognition and offers a wide array of possible avenues to explore. First, an in-depth exploration of the various properties that non-lexical auditory stimuli may have is certainly warranted. Indeed, future studies could focus on systematically manipulating the auditory stimuli with varying degrees of semantic and lexical status. For example, rather than use common, every day sounds, it would be interesting to see if processing for non-words, compared to words, would yield different results. Arguably, non-words have no semantic connection though they maintain lexical properties. This manipulation would allow for the isolation of semantic content between words and non-words, which would help to elucidate the role of semantic access in processing for auditory stimuli.

Alternatively, melodic perception could also be explored. Auditory patterns present in music are quite ubiquitous in daily life, and are often presented under conditions in which they would be ignored, at least to some extent (e.g., elevators, waiting rooms, grocery stores, etc.). While lyrics contain lexical and semantic properties, it is probable that a novel melodic sequence does not. This would allow for the comparison between purely lexical auditory information (i.e., words) and perhaps purely non-lexical auditory information (i.e., novel melodies). Again, preliminary work would first require determining the extent to which melodic perception may be tied to semantic or lexical representations, but this would be an interesting direction nonetheless.

Additional consideration may also be given to the timing and spatial location of stimulus presentation. Recall that facilitation by temporal-alignment is thought to proceed mainly on a temporal dimension (Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe, 2001), meaning that facilitation may operate in a manner similar to that of the attentional blink phenomenon (Chun & Potter, 1995; Shapiro, 1994; Wolfe & Gray, 2007).

However, it should be noted that the current experiments presented all attended and ignored information with temporal and spatial alignment. Therefore, it is not possible to be determine whether it was the temporal or spatial coincidence that primarily drives the observed facilitatory effect of temporal-alignment. Indeed, previous research has demonstrated that, under cross-modal conditions, visual information must precede auditory information in order for an agent to perceive the two items as appearing simultaneously in a TOJ task (Zampini, Shore, & Spence, 2005), suggesting that ideal conditions for signal integration may not be simultaneous onset. Thus, in order to disentangle the contributions of temporal and spatial alignment, future projects may consider systematically manipulating the timing and spatial location of stimulus presentation during the primary task in an effort to determine which factor is driving the facilitatory effect.

Another avenue worth investigating centers around disambiguating the extent to which the observed facilitatory effects are a result of enhanced encoding or retrieval. While an earlier investigation using this same paradigm to investigate inattentional blindness suggests that inhibitory mechanisms take place during encoding (Rees, et al., 1999), facilitation and recall have not been explicitly studied with this task. One method that may be employed to investigate the nature of the recognition rates during the surprise recognition test would be to include a subjective measure of confidence in participant's old/new judgments. Specifically, in the surprise recognition task, participants could be given a likert scale on their confidence level in order to explore this measure across target-aligned, non-aligned, and foil items. Alternatively, participants could be surveyed on the nature of their decision in a manner akin to those utilized by Roediger and McDermott (1995) wherein participants were asked to state whether or not recalled items were "remembered" or "known" to be present in a previously heard list of items. Here, "remember" implies that the participant has an explicit memory of the item occurring in the list while "know" implies that an item appeared to be familiar but the participant had no explicit memory for the item itself. In either case, higher confidence ratings or higher "remember" reports for target-aligned items would suggest that participants are experiencing facilitated retrieval compared to non-aligned items.

It would also be worthwhile to explore the extent to which processing for these stimulus types across the visual and auditory sensory modalities may be *inhibited* rather than facilitated. Indeed, extensive research using the same paradigm employed in this dissertation has demonstrated inhibited processing for previously ignored items when the presentation rate and frequency of target-alignment are reduced (Dewald et al., 2011; Dewald & Sinnett, 2011; Walker, Ciraolo, Dewald, & Sinnett, 2016), though how inhibition may proceed for different stimulus types or across different sensory modalities remains under explored.

Finally, examining these properties across the life span or in special populations could provide useful information regarding the developmental trajectory of the cognitive mechanisms involved in the facilitated (and inhibited) processing for task-irrelevant stimuli. For example, a cross-sectional comparison between healthy young adults, middle-aged adults, and elderly individuals would provide groundwork for a model of cognitive aging focused around the interplay between sensory and attentional capabilities. This could also be taken a step further to include special populations such as those chronically infected with HIV, as this disease has been demonstrated to impart cognitive deficits even among younger adults (Elbirt, Mahlab-Guri, Bazalel-Rosenberg, Attali, & Asher, 2015; MacArthur & Brew, 2010).

Taken together, this project represents a rich and fruitful vein of study with the potential to greatly expand our understanding of the human cognitive system. Given the paucity of research focusing explicitly on how we process the various forms of stimuli that are present in daily life, and how this may be augmented by the sensory modality in which these stimuli are presented, one may conclude that this investigation remains in its infancy. Thus, the potential for continued exploration is wide and promising!

Appendix A: IRB Approval Letter

UNIVERSITY of HAWAI'I®
SYSTEM

Office of Research Compliance
Human Studies Program

MEMORANDUM
CR

January 29, 2019

TO: Scott Sinnett, Ph,D.
Principal Investigator
Psychology Department

FROM: Victoria Rivera Victoria Rivera
Interim Director

SUBJECT: CHS#21455 "Information Processing Within and Across Sensory Modalities"

This is to acknowledge receipt of your response received January 14, 2019 to the stipulations issued by the Human Studies Program during its review of the project identified above at its meeting on January 11, 2019. The information you provided satisfactorily addressed the Human Studies Program stipulations, and the project is approved for one year, effective January 29, 2019.

This memorandum is your record of the Human Studies Program approval of this study. Please maintain it with your study records.

The Human Studies Program approval for this project will expire on January 28, 2020. If you expect your project to continue beyond this date, you must submit an application for renewal of this Human Studies Program approval. The Human Studies Program approval must be maintained for the entire term of your project.

If, during the course of your project, you intend to make changes to this study, you must obtain approval from the Human Studies Program prior to implementing any changes. If an Unanticipated Problem occurs during the course of the study, you must notify the Human Studies Program within 24 hours of knowledge of the problem. A formal report must be submitted to the Human Studies Program within 10 days. The definition of "Unanticipated Problem" may be found at: https://www.hawaii.edu/researchcompliance/policies-guidance, and the report form may be downloaded here: https://www.hawaii.edu/researchcompliance/report-protocol-violation-or-unanticipated-problem.

You are required to maintain complete records pertaining to the use of humans as participants in your research. This includes all information or materials conveyed to and received from participants as well as signed consent forms, data, analyses, and results. These records must be

2425 Campus Road, Sinclair 10
Honolulu, Hawai'i 96822
Telephone: (808) 956-5007 • Fax: (808) 956-9150
An Equal Opportunity/Affirmative Action Institution

Full Accreditation

maintained for at least three years following project completion or termination, and they are subject to inspection and review by the Human Studies Program and other authorized agencies.

Please notify this office when your project is completed. Upon notification, we will close our files pertaining to your project. Reactivation of the Human Studies Program approval will require a new Human Studies Program application.

Please contact this office if you have any questions or require assistance. We appreciate your cooperation, and wish you success with your research.

Appendix B: Consent Form

University of Hawai'i
Consent to Participate in a Research Project
Scott Sinnett, Principal Investigator
Maegen Walker, Graduate Student

Project title: *Information processing within and across sensory modalities*

Aloha! My name is Maegen Walker and you are invited to take part in a research study. I am a graduate student at the University of Hawai'i at Mānoa in the Department of Psychology. As part of the requirements for earning my graduate degree, I am doing a research project.

What am I being asked to do?
If you participate in this project, you will be asked to respond to various items, such as pictures, words, or sounds, presented to you via a computer.

Taking part in this study is your choice.
Your participation in this project is completely voluntary. You may stop participating at any time. If you stop being in the study, there will be no penalty or loss to you. Your choice to participate or not participate will not affect your rights to services at the UH Career Development and Counseling Program.

Why is this study being done?
This study aims to gain a deeper understanding of how humans direct attention and perceive auditory and visual stimuli in our environments. One of the goals of this project is to assess how our ability to attend to and perceive stimuli is altered through the combination of two or more sensory experiences of various stimuli when compared to experiencing those stimuli with just one of our senses. We are also interested in determining how our ability to pay attention and process information changes as we age. We hope to develop broader knowledge of the differences between youths and older populations with regard to these abilities in an effort to contribute to an area of science that has yet been left relatively unexplored. Developing our knowledge of how our attention and perception changes with age may shed light on various mental disabilities associated with natural aging.

What will happen if I decide to take part in this study?
If you agree to participate, the experiment will take about 30-60 minutes of your time. You will spend most of this time seated in front of a computer monitor. You will be presented with a visual stream of objects originating from the screen and/or auditory sounds or spoken words originating from the speakers placed beside the screen or a pair of headphones. You will be able to adjust the volume of the auditory sounds to a level of your comfort. You will be required to respond to specific targets that occur in the auditory or visual stream by pressing different keys on the keyboard, or a foot pedal under your foot.

Before beginning the experiment, you will receive ample instruction and training on the task. If you are not sure about any instructions, or wish to have more practice, do not hesitate to ask. Throughout the experiment, you will be given ample opportunity to take breaks, should you wish, and may discontinue your participation at any time without loss of compensation of penalty. While you are engaging in this task a remote, non-invasive camera called an eye-tracker will monitor your eye gaze. This tool is designed to monitor the direction of your eye gaze only and will not collect any additional or identifying information about you.

Upon completion of the study you may be asked to take part in a brief survey, which will take approximately five (5) minutes to complete.

University of Hawai'i
Consent to Participate in a Research Project
Scott Sinnett, Principal Investigator
Maegen Walker, Graduate Student

Project title: *Information processing within and across sensory modalities*

What are the risks and benefits of taking part in this study?
I believe there is little risk to you for participating in this research project. The minimum risk includes feelings of mild exhaustion, dizziness or disorientation from viewing the images presented to you on the screen. To minimize these risks we will give you ample opportunity to take breaks throughout your participation for the amount of time you feel it is necessary to be able to continue. You will be continuously monitored during your participation and if at any time you begin to exhibit any signs of discomfort we will discontinue the experiment immediately. Termination of the experiment as a result of any discomfort you might experience will in no way effect your compensation.

If you suffer form any pre-existing conditions or risk factors (epilepsy for example) that you feel would prevent you from safely participating in this study, please do not participate. Your non-participation as a result of a pre-existing condition or risk factor will in no way effect your compensation.

There are no direct benefits to you, however your participation will contribute to the enhancement of our understanding of human cognition and how our mental abilities change over time. The information collected here will be compared to an aging population and this comparison may be used to enhance our knowledge of mental decline with age and contribute to a greater understanding of human mental capabilities in general.

Privacy and Confidentiality:
Your identity will be kept strictly confidential. All documents will be identified only by a subject code number and kept in a locked filing cabinet. You will not be identified by name in any reports of the completed study. Data that will be kept on a computer hard disk will also be identified only by your subject code number and will be password protected so that only the principle investigator, Dr. Scott Sinnett, his graduate students, and research assistants will have access to it. Other agencies that have legal permission have the right to review research records. The University of Hawai'i Human Studies Program has the right to review research records for this study. Following the completion of the study, the data will be transferred to a CD and stored in a locked filing cabinet. Note, the results of this study will be used to write a scientific report.

Compensation:
You will receive a maximum of two (2) SONA credits for your voluntary participation in this study.

Future Research Studies:
The data may be used for future research studies or distributed to another investigator for future research studies and we will not seek further approval from you for these future studies.

Questions:
If you have any questions about this study, please email me at [maegenw@hawaii.edu]. You may also contact my advisor, Dr. Scott Sinnett, at [808-956-6272 & ssinnett@hawaii.edu] to discuss

problems, concerns and questions; obtain information; or offer input with an informed individual who is unaffiliated with the specific research protocol.

Please visit http://go.hawaii.edu/jRd for more information on your rights as a research participant.

If you agree to participate in this project, please sign and date this signature page and return it to the experimenter.

Keep a copy of the informed consent for your records and reference.

Signature(s) for Consent:

I give permission to join the research project entitled, "*Information processing within and across sensory modalities"*

Name of Participant (Print): ____

Participant's Signature: ______________________________________

Date: ________________________

Mahalo!

Appendix C: Debriefing Form

The University of Hawai'i at Manoa

Debriefing
Attention and Perception

Dear Participant,

During this study, you were asked to participate in one or more activities via computer, which included auditory, visual, or text displayed on the computer. You were asked to respond to a stream of objects by pressing different keys on the keyboard, or a foot pedal under your foot. You'll notice that we did not tell you that you were going to be asked to identify some of the objects we told you to ignore during the first part of the experiment. We did not tell you about this part of the experiment because if you had known that we were going to ask you about the objects later the results of the study would be compromised. It is important that future participants are also unaware of this aspect of the study so please keep this information confidential and do not discuss the nature of the experiment with anyone.

You are reminded that your original consent document included the following information: Your participation in this study is entirely voluntary and you may refuse to participate or withdraw from the study at any time without penalty.

If you have questions about your participation in the study, please contact me at maegenw@hawaii.edu or my faculty advisor, Dr. Sinnett at ssinnett@hawaii.edu

If you have questions about your rights as a research participant, you may contact the UH Human Studies Program, by phone at (808) 956-5007, or University of Hawai‘i Institutional Review Board at uhirb@hawaii.edu.

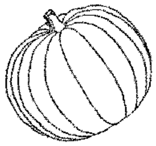
Please again accept our appreciation for your time, participation, and cooperation.

Sincerely,

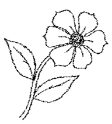
Maegen Walker

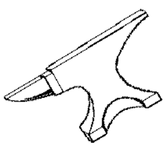
Appendix D: List of Stimuli

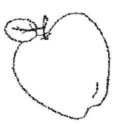
Pictures

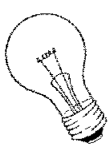

1.

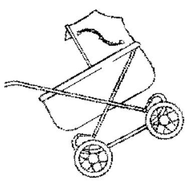

2.

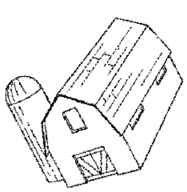

3.

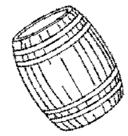

4.


5.


6.

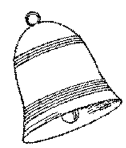

7.

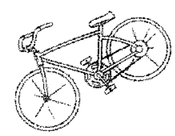

8.

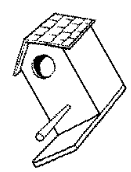

9.


10.

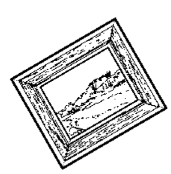

11.

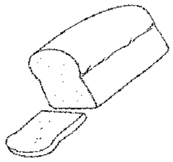

12.


13.


14.

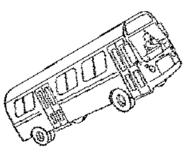

15.

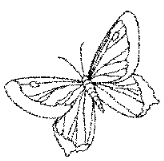

16.

17. 

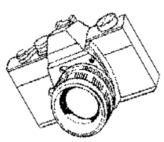
18. 

19. 

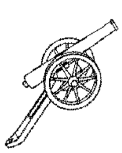
20. 

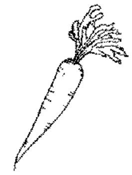
21. 

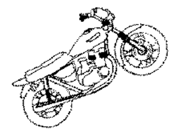
22. 

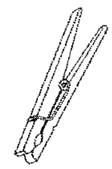
23. 

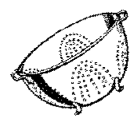
24. 

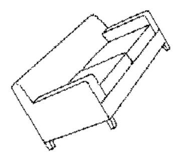
25. 

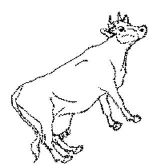
26. 

27. 

28. 

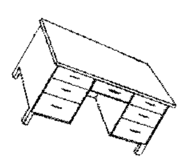
29. 

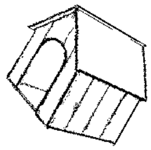
30. 

31. 

32. 

33. 

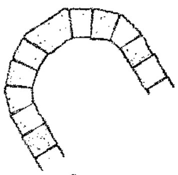
34. 



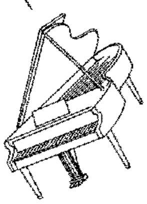
35.

36.

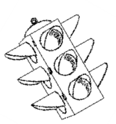
37.



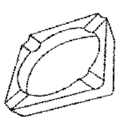
38.



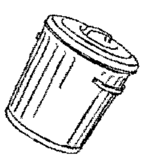
39.



40.

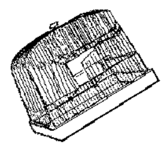


41.

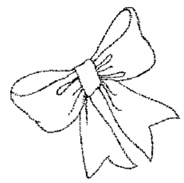


42.

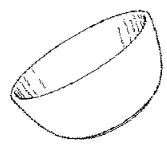
43.



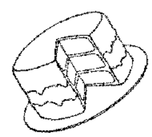
44.



45.

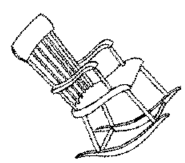


46.

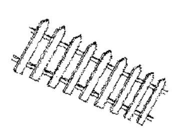


47.

48.



49.



50.

Sounds

1. Gunshot
2. Lighter Flick
3. Fly Buzz
4. Low Frequency Bell
5. Sharp Whistle
6. Burp
7. Stapler
8. Cat Meow
9. Droop Whistle
10. Cell Phone Ring
11. High Pitch Cricket
12. Crunch
13. Crow Caw
14. Dog Bark
15. Chitter
16. Doorbell
17. Water Droplet
18. Frog Ribbit
19. Door Knock
20. Machine Gun
21. Telephone Ring
22. Pool Ball Strike
23. Old Car Start
24. Pouring Liquid
25. Bilabial Trill
26. Traffic Honk
27. Typewriter
28. Clapping
29. Tire Screech
30. Cash Register
31. Cat Growl
32. Mid Frequency Bell
33. Reveille Horn
34. Low Pitch Cricket
35. Fart
36. Window Shatter
37. High Frequency Bell
38. Laser
39. Siren
40. Rooster
41. School Bell
42. Cow Moo
43. Crash
44. Ah-Choo
45. Choir
46. Truck Horn
47. Foghorn
48. Car Horn
49. Pip
50. Bird Chirp

Words

1. Trail
2. Still
3. Result
4. Public
5. Horn
6. Land
7. Volume
8. Thing
9. Tube
10. River
11. Data
12. Future
13. Change
14. Dark
15. Family
16. Modern
17. Line
18. Wall
19. Blue
20. Debate
21. Court
22. Game
23. Wife
24. County
25. Third
26. Read
27. Voice
28. Name
29. Music
30. Fire
31. Common
32. Level
33. Night
34. Type
35. Secret
36. Case
37. Work
38. Remote
39. Week
40. Well
41. Table
42. Love
43. Church
44. Peace
45. Army
46. Party
47. Girl
48. Time
49. Done
50. Paper

References

Adaval, R., Saluja, G., & Jiang, Y. (2019). Seeing and thinking in pictures: A review of visual information processing. *Consumer Psychology Review*, *2*(1), 50-69.

Agresti, A. (2002). Categorical Data Analysis. John Wiley & Sons. *Inc., Publication*.

Agresti, A., & Kateri, M. (2011). *Categorical data analysis* (pp. 206-208). Springer Berlin Heidelberg.

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343-348.

Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General*, *138*(3), 400-415.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367-10371.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*(2), 82-91.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation*, *8*, 47-89. Academic press.

Baddeley, A. D., & Weiskrantz, L. E. (1993). *Attention: Selection, awareness, and control: A tribute to Donald Broadbent*. Clarendon Press/Oxford University Press.

Barber, P. J. (2015). *Applied cognitive psychology: An information-processing framework*. Routledge.

Ballas, J. A., & Howard Jr, J. H. (1987). Interpreting the language of environmental sounds. *Environment and Behavior*, *19*(1), 91-114.

Ballesteros, S. (2014). *Cognitive approaches to human perception*. Psychology Press.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In *Advances in psychology*. 129, 347-362.

Besner, D., Risko, E. F., & Sklair, N. (2005). Spatial attention as a necessary preliminary to early processes in reading. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *59*(2), 99-108.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*(5), 512-528.

Bles. M., & Jansma, B.M. (2008). Phonological processing of ignored distractor pictures, an fMRI investigation. *BMC Neuroscience.* 9(1), 20. doi: 10.1186/1471-2202-9-20

Bransford, J. (1979). Human cognition: Learning, understanding, and remembering. Belmont, CA: Wadsworth.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, *64*(3), 205-215.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.

Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics*, *33*(5), 487-493.

Carr, T.H, McCauley, C., Sperber, R.D., & Parmelee, C.M., (1982). Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance. 8(6),* 757-777.

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, *102*(3), 321-340.

Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*. John Wiley & Sons.

Chatterjee, S., Hadi, A. S., & Price, B. (2000). Regression analysis by example John Wiley & Sons. *Inc., New York*.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409-412.

Colavita, F. B., & Weisberg, D. (1979). A further investigation of visual dominance. *Perception & Psychophysics*, *25*(4), 345-347.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 226–234.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331-335.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*, *124*(3), 367-372.

de Zubicaray, G., McMahon, K., Eastburn, M., Pringle, A., & Lorenz, L. (2006) Classic identity negative priming involves accessing semantic representations in the left anterior temporal cortex. *Neuroimage*. *33(1)*, 383-390.

Deacon, D., Hewitt, S., Yang, C.M., & Nagata, M. (2000). Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. *Cognitive Brain Research, 9(2),* 137-146.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., … & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395(6702*), 597-600.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80-90.

Dewald, A.D., & Sinnett, S. (2012). Enhanced performance for recognition of irrelevant target-aligned auditory stimuli: Unimodal and cross-modal considerations. *In Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Societ*y, 294-299.

Dewald, A. D., Sinnett, S., & Doumas, L. A. (2011). Conditions of directed attention inhibit recognition performance for explicitly presented target-aligned irrelevant stimuli. *Acta Psychologica*, *138*(1), 60-67.

Dewald, A. D., Sinnett, S., & Doumas, L. A. (2013). A window of perception when diverting attention? Enhancing recognition for explicitly presented, unattended, and irrelevant stimuli by target alignment. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1304-1312.

Dodd, M. D., & Pratt, J. (2007). Rapid onset and long-term inhibition of return in the multiple cuing paradigm. *Psychological Research*, *71*(5), 576-582.

Dodd, M. D., Van der Stigchel, S., & Hollingworth, A. (2009). Novelty is not always the best policy: Inhibition of return and facilitation of return as a function of visual task. *Psychological Science*, *20*(3), 333-339.

Donohue, S. E., Roberts, K. C., Grent, T., & Woldorff, M. G. (2011). The cross-modal spread of attention reveals differential constraints for the temporal and spatial linking of visual and auditory stimulus events. *Journal of Neuroscience*, *31*(22), 7982-7990.

Driver, J., & Spence, C. (1998). Attention and the crossmodal structure of space. *Trends in Cognitive Sciences*, *2*, 254– 262.

Driver, J., & Spence, C. (2004). Crossmodal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), Crossmodal space and crossmodal attention. Oxford, UK: Oxford University Press.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272-300.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387, 808–810.

Edell, J. A., & Staelin, R. (1983). The information processing of pictures in print advertisements. *Journal of Consumer Research*, *10*(1), 45-61.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784-1790.

Elbirt, M. D., Mahlab-Guri, K., Bazalel-Rosenberg, S., Attali, M., & Asher, I. (2015). HIV-associated neurocognitive disorders (HAND). *The Israel Medical Association Journal, 17(1),* 54-59.

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers,* 35(1), 116-124.

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. *Psychological Bulletin*, *81*(5), 284-310.

Fry, D.B. (1968). Prosodic phenomena. In Manual of Phonetics (ed. B. Malmberg). Amsterdam: North Holland.

Gibbs, R., Davies, G., & Chou, S. (2016). A systematic review on factors affecting the likelihood of change blindness. *Crime Psychology Review*, *2*(1), 1-21.

Gibson, E.J. (1969). Principles of perceptual learning and development. East Norwalk, CT, US: Appleton-Century-Crofts.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350-363.

Giraudet, L., St-Louis, M. E., Scannella, S., & Causse, M. (2015). P300 event-related potential as an indicator of inattentional deafness?. *PLoS one*, *10*(2), e0118556.

Harrell, F. E. Jr., (2001). Regression modeling strategies. New York: Springer.

Harrell, F. E. Jr.,(2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. New York: Springer.

Haykin, S., & Chen, Z. (2005). The cocktail party problem. *Neural Computation*, *17*(9), 1875-1902.

Hazan, V. & Rosen, S. (1991) Individual variability in the perception of cues to place contrasts in initial stops. Percept. *Psychophysics*. *49*, 187-200.

Henny Yeung, H., Bhatara, A., & Nazzi, T. (2018). Learning a Phonological Contrast Modulates the Auditory Grouping of Rhythm. *Cognitive Science*, *42*(6), 2000-2020.

Hitch, G., & Baddeley, A. D. (2017). Working memory. In *Exploring Working Memory* (pp. 43-79). Routledge.

Hogaboam, T.W., Pellegrino, J.W., (1978). Hunting for individual differences in cognitive processes: Verbal ability and semantic processing of pictures and words. *Memory & Cognition. 6*(2), 189-193.

Huitt, W. (2003). The information processing approach to cognition. *Educational Psychology Interactive*, *3*(2), 53-60.

Humphreys, G. W. (1981). On varying the span of visual attention: Evidence for two modes of spatial attention. *The Quarterly Journal of Experimental Psychology*, *33*(1), 17-30.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434-446.

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *Journal of Experimental Psychology: General*, *138*(2), 291-306.

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(2), 110-126.

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effect in verbal memory. *Psychological Bulletin*, *85*(3), 532-554.

Kim, R. S., Seitz, A. R., & Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One*, *3*(1), e1532.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, *55*(4), 352-358.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16-22.

Koreimann, S., Gula, B., & Vitouch, O. (2014). Inattentional deafness in music. *Psychological Research*, *78*(3), 304-312.

LaBerge, D. (2014). Perceptual learning and attention. *Learning and Cognitive Processes*, *4*, 237-273.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293-323.

Lachman, R., Lachman, J. L., & Butterfield, E. C. (2015). *Cognitive Psychology and Information Processing: An Introduction*. Psychology Press.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): still no identification without attention. *Psychological Review*, *111*(4), 880-913.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, *27*(8), 1155-1163.

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*(4), 405-414.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451-468.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current directions in psychological science*, *19*(3), 143-148.

Lehiste, I. (1970). Suprasegmentals. Cambridge, Massachusetts: M IT Press.

Liu, T., Abrams, J., & Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science*, *20*(3), 354-362.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740-749.

Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1–36.

Mack, A., & Rock, I. (1998a). *Inattentional blindness* (Vol. 33). Cambridge, MA: MIT press.

Mack, A., & Rock, I. (1998b). Inattentional blindness: Perception without attention. *Visual Attention*, *8*, 55-76.

McArthur, J. C., & Brew, B. J. (2010). HIV-associated neurocognitive disorders: is there a hidden epidemic?. *Aids*, *24*(9), 1367-1370.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*(2), 163-203.

Marcel, A. J. (1980). Conscious and pre-conscious recognition of polysemous wards. In R. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Massaro, D. W. (2014). Auditory information processing. *Handbook of Learning and Cognitive Processes (Volume 4): Attention and Memory*, 275-320.

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73*(6), 1780-1789.

McClelland, J. L. (2000). Connectionist models of memory. *The Oxford Handbook of Memory*, 583-596.

McClelland, J. L., & Elman, J. L. (1986). The Trace model of speech perception. *Cognitive Psychology*, 18, 1–86.

Mcdonald, J. J., Teder-Sälejärvi, W. A., Russo, F. D., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, *15*(1), 10-19.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746-748.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual–auditory object recognition in humans: a high-density electrical mapping study. *Cerebral Cortex*, *14*(4), 452-465.

Molloy, K., Lavie, N., & Chait, M. (2019). Auditory figure-ground segregation is impaired by high visual load. *Journal of Neuroscience*, *39*(9), 1699-1708.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 339-352.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154-163.

Morey, C. C. (2018). Perceptual grouping boosts visual working memory capacity and reduces effort during retention. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12355

Murphy, G., & Greene, C. M. (2015). High perceptual load causes inattentional blindness and deafness in drivers. *Visual Cognition*, *23*(7), 810-814.

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research*, *352*, 40-48.

Neill, W. T., Lissner, L. S., & Beck, J. L. (1990). Negative priming in same-different matching: Further evidence for a central locus of inhibition. *Perception & Psychophysics*, *48*(4), 398-400.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*(4), 480-494.

Ngo, M. K., Cadieux, M. L., Sinnett, S., Soto-Faraco, S., & Spence, C. (2011). Reversing the Colavita visual dominance effect. *Experimental brain research*, *214*(4), 607-618.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189–234.

Notter, M. P., Hanke, M., Murray, M. M., & Geiser, E. (2018). Encoding of Auditory Temporal Gestalt in the Human Brain. *Cerebral Cortex*, *29*(2), 475-484.

Nousak, J. M. K., Deacon, D., Ritter, W., & Vaughan Jr, H. G. (1996). Storage of information in transient auditory memory. *Cognitive Brain Research*, *4*(4), 305-317.

Parker, J. L., & Robinson, C. W. (2018). Changes in multisensory integration across the life span. *Psychology and Aging*, *33*(3), 545.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688-691.

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision research*, *45*(14), 1867-1875.

Poort, J., Raudies, F., Wannig, A., Lamme, V. A., Neumann, H., & Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron*, *75*(1), 143-156.

Posner, M. I. (1978). *Chronometric explorations of mind*. Oxford, England: Lawrence Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychological Review*, *83*(2), 157-171.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160-174.

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, *77*(2), 483-492.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science*, *286*(5449), 2504-2507.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1303-1323.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803-814.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society London B.*, *336*(1278), 367-373.

Rosen, S. & Fourcin, A. J. (1986). Frequency selectivity and the perception of speech. In Frequency selectivity in hearing (ed. B. C. J. Moore), pp. 373-487. London: Academic Press.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1987). *Parallel Distributed Processing* (Vol. 1). Cambridge, MA: MIT press.

Ruz, M., Wolmetz, M. E., Tudela, P., & McCandliss, B. D. (2005). Two brain pathways for attended and ignored words. *Neuroimage*, *27*(4), 852-861.

Ruz, M., Worden, M. S., Tudela, P., & McCandliss, B. D. (2005). Inattentional amnesia to words in a high attentional load task. *Journal of Cognitive Neuroscience*. 17(5), 768-776.

Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics*, *54*(3), 406-416.

Sasaki, Y., Náñez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*(1), 53-60.

Seitz, A. R., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive?. *Nature*, *422*(6927), 36.

Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*(7), 329-334.

Seitz, A. R., & Watanabe, T. (2008). Is task-irrelevant learning really task-irrelevant?. *PloS one*, *3*(11), e3792.

Seitz, A. R., & Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision Research*, *49*(21), 2604-2610.

Self, M. W., van Kerkoerle, T., Super, H., & Roelfsema, P. R. (2013). Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Current Biology*, *23*(21), 2121-2129.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*(6814), 788.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411-417.

Shapiro, K. L. (1994). The attentional blink: The brain's “eyeblink”. *Current Directions in Psychological Science*, *3*(3), 86-89.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, *76*(2), 179-193.

Shipp, S., & Zeki, S. (1985). Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. *Nature*, *315*(6017), 322-324.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*(3), 205-212.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology*, *30*(1), 363-396.

Simon, J., & Winkler, I. (2018). The role of temporal integration in auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(11), 1683-1693.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059-1074.

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition*, *11*(1), 78-97.

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *The Quarterly Journal of Experimental Psychology*. 59(8), 1425-1442.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics*, *69*(5), 673-686.

Skóra, Z., & Wierzchoń, M. (2016). The level of subjective visibility at different stages of memory processing. *Journal of Cognitive Psychology*, *28*(8), 965-976.

Smith, T. J., & Henderson, J. M. (2009). Facilitation of return during scene viewing. *Visual Cognition*, *17*(6-7), 1083-1108.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*. *109*(4), 373-392.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174-215.

Snyder, J. S. (2015). Sound perception: Rhythmic brain activity really is important for auditory segregation. *Current Biology*, *25*(24), R1173-R1175.

Sørensen, T. A. (2017). On the relationship between short-and long-term memory. In *Neuroscience Day*.

Soto-Faraco, S., Morein-Zamir, S., & Kingstone, A. (2005). On audiovisual spatial synergy: The fragility of the phenomenon. *Perception & Psychophysics*, *67*(3), 444-457.

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research*, *176*, 245-258.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005-1030.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1-22.

Spence, C., Parise, C., & Chen, Y. C. (2012). The Colavita visual dominance effect. In: Murray MM, Wallace MT, editors. The Neural Bases of Multisensory Processes. Boca Raton (FL): CRC Press/Taylor & Francis.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799-832.

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 539-550.

Stein, B. E., Stanford, T. R., Wallace, M. T., Vaughan, J. W., & Jiang, W. (2004). Crossmodal spatial interactions in subcortical and cortical circuits. *Crossmodal Space and Crossmodal Attention*, 25-50.

Stevens, K. N. (1980). Acoustic correlates of some phonetic categories. *The Journal of the Acoustical Society of America*, *68*(3), 836-842.

Stevens, K. N. (1981). Constraints imposed by the auditory system on the properties used to classify speech sounds: Data from phonology, acoustics, and psychoacoustics. In *Advances in Psychology. 7*, 61-74.

Stevens, K. N., & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. *Perspectives on the Study of Speech*, 1-38.

Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences*, *106*(52), 22456-22461.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662.

Swallow, K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition*, *115*(1), 118-132.

Swallow, K. M., & Jiang, Y. V. (2011). The role of timing in the attentional boost effect. *Attention, Perception, & Psychophysics*, *73*(2), 389-404.

Teki, S., Barascud, N., Picard, S., Payne, C., Griffiths, T. D., & Chait, M. (2016). Neural correlates of auditory figure-ground segregation based on temporal coherence. *Cerebral Cortex*, *26*(9), 3669-3680.

Teng, X., Tian, X., Rowland, J., & Poeppel, D. (2017). Concurrent temporal channels for auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. *PLoS Biology*, *15*(11), e2000812.

Theunissen, F. E., & Elie, J. E. (2014). Neural processing of natural sounds. *Nature Reviews Neuroscience*, *15*(6), 355-366.

Thurlow, W. R., & Jack, C. E. (1973). Certain determinants of the “ventriloquism effect”. *Perceptual and Motor Skills*, *36*(3), 1171-1184.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology*. 37(4), 571-590.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition*, *16*(1), 64-70.

Tipper, S. P., MacQueen, G. M., & Brehaut, J. C. (1988). Negative priming between response modalities: Evidence for the central locus of inhibition in selective attention. *Perception & Psychophysics*, *43*(1), 45-52.

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, *83*(2), 93-154.

Tillmann, J., & Swettenham, J. (2017). Visual Perceptual Load Reduces Auditory Detection in Typically Developing Individuals but Not in Individuals With Autism Spectrum Disorders. *Neuropsychology*, *31*(2), 181-190.

Tóth, B., Kocsis, Z., Háden, G. P., Szerafin, Á., Shinn-Cunningham, B. G., & Winkler, I. (2016). EEG signatures accompanying auditory figure-ground segregation. *Neuroimage*, *141*, 108-119.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282-299.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194-214.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136.

Tsushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science*, *314*(5806), 1786-1788.

Tsushima, Y., Seitz, A. R., & Watanabe, T. (2008). Task-irrelevant learning occurs only when the irrelevant feature is weak. *Current Biology*, *18*(12), R516-R517.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254-256.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325-329.

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*, 75-94.

Yantis, S. (1996). 2. Attentional capture in vision. In *Converging Operations in the Study of Visual Selective Attention*. 45-76.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: evidence from equiluminant visual objects. *Journal of experimental psychology: Human perception and performance*, *20*(1), 95.

Waechter, S., Besner, D., & Stolz, J. A. (2011). Basic processes in reading: Spatial attention as a necessary preliminary to orthographic and semantic processing. *Visual Cognition*, *19*(2), 171-202.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172-1217.

Walker, M., Ciraolo, M., Dewald, A., & Sinnett, S. (2017). Differential processing for actively ignored pictures and words. *PloS One*, *12*(1), e0170520.

Walker, M., Dewald, A., & Sinnett, S. (2014). The role of modality congruence in the presentation and recognition of task-irrelevant stimuli in dual task paradigms. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36), 1736-1741.

Wang, Y., Zhang, J., Zou, J., Luo, H., & Ding, N. (2018). Prior knowledge guides speech segregation in human auditory cortex. *Cerebral Cortex, 29(4),* 1561-1571, https://doi.org/10.1093/cercor/bhy052

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, *413*(6858), 844-848.

Watanabe, T., & Sasaki, Y. (2015). Perceptual learning: toward a comprehensive theory. *Annual Review of Psychology*, *66*, 197-221.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), Varieties of attention (pp. 63 – 101). Orlando, FL: Academic Press.

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*. *20*(1), 6-10.

Wingfield, A. (2016). Evolution of models of working memory and cognitive resources. *Ear and Hearing*, *37*, 35S-43S.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202-238.

Wolfe, J. M. (2001). Guided Search 4.0: A guided search model that does not require memory for rejected distractors. *Journal of Vision*, *1*(3), 349-349.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419-433.

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. *Integrated Models of Cognitive Systems*, 99-119.

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience letters*, *381*(3), 217-222.